(NASA-CR-173010) THE UARS AND OPEN DATA CONCEPT AND ANALYSIS STUDY Final Report (Engineering and Economics Research, Inc.) 125 p HC A06/MF A01 CSCL 05B

N83-34846

Unclas
G3/82 36087

**ENGINEERING & ECONOMICS RESEARCH, INC.**

# UARS AND OPEN DATA SYSTEM CONCEPT AND ANALYSIS STUDY

FINAL REPORT

Prepared for:

National Aeronautics and Space Administration
Goddard Space Flight Center
Greenbelt, Maryland 20771

Under Contract Number: NAS5-26962


Submitted by:

M. Mittal
J. Nebb
H. Woodward

Engineering and Economics Research, Inc.
1951 Kidwell Drive
Vienna, Virginia 22180


March, 1983

TABLE OF CONTENTS

TABLE OF CONTENTS (Concluded)

LIST OF FIGURES

LIST OF TABLES

LIST OF TABLES (Concluded)

# 1.0 INTRODUCTION

## 1.0 INTRODUCTION

This report offers alternative concepts for a common design for the UARS and OPEN Central Data Handling Facility (CDHF) (see Section 4). The designs are consistent with requirements shared by UARS and OPEN (Section 2 and Appendix 2) and the data storage and data processing demands of these missions (Section 3). Because more detailed information is available for UARS, the design approach has been to size the system and to select components for a UARS CDHF, but in a manner that does not optimize the CDHF at the expense of OPEN. Costs for alternative implementations of the UARS designs are presented in Sections 4.1 and 4.2, showing that the system design does not restrict the implementation to a single manufacturer. Processing demands on the alternate UARS CDHF implementations are then discussed in Section 4.3. With this information at hand together with estimates for OPEN processing demands (Section 3.2.2), it is shown that any shortfall in system capability for OPEN support can be remedied by either component upgrades or array processing attachments rather than a system redesign.

In addition to a common system design, it is shown in Section 5 that there is significant potential for common software design, especially in the areas of data management software and non-user-unique production software.

The report then discusses archiving the CDHF data (Section 6). Following that, cost examples for several modes of communications between the CDHF and Remote User Facilities are presented (Section 7).

The report concludes with a discussion of the potential application of technologies expected to reach fruition before the mission timeframe (Section 8).

# 2.0 UARS AND OPEN SYSTEM LEVEL REQUIREMENTS

## 2.0 UARS AND OPEN SYSTEM LEVEL REQUIREMENTS

Based upon available documentation (see Bibliography) and input from GSFC technical personnel, a list of OPEN and UARS missions, system level requirements, assumptions and intercomparisons was generated (see Appendix A), in which particular emphasis was placed upon the Central Data Handling Facility (CDHF). It is seen that there are a number of system level functions common to both a UARS and an OPEN CDHF. The major of these common functions are:

- Data ingest of playback data
- Routine production processing of the data
- Data management
- Investigator communications.

The distribution of these functions within the proposed CDHF concepts is defined in Section 4.0.

### 2.1 UARS and OPEN Systems Elements

Not only do the UARS and OPEN CDHF's share similar system level requirements, but their relations to institutional facilities are also similar. This is illustrated in Figure 2.1-1. As shown in the figure, in addition to the CDHF, both the UARS and OPEN ground systems consist of the following functional elements:

- Data capture
- Orbit determination
- Attitude determination
- Command management

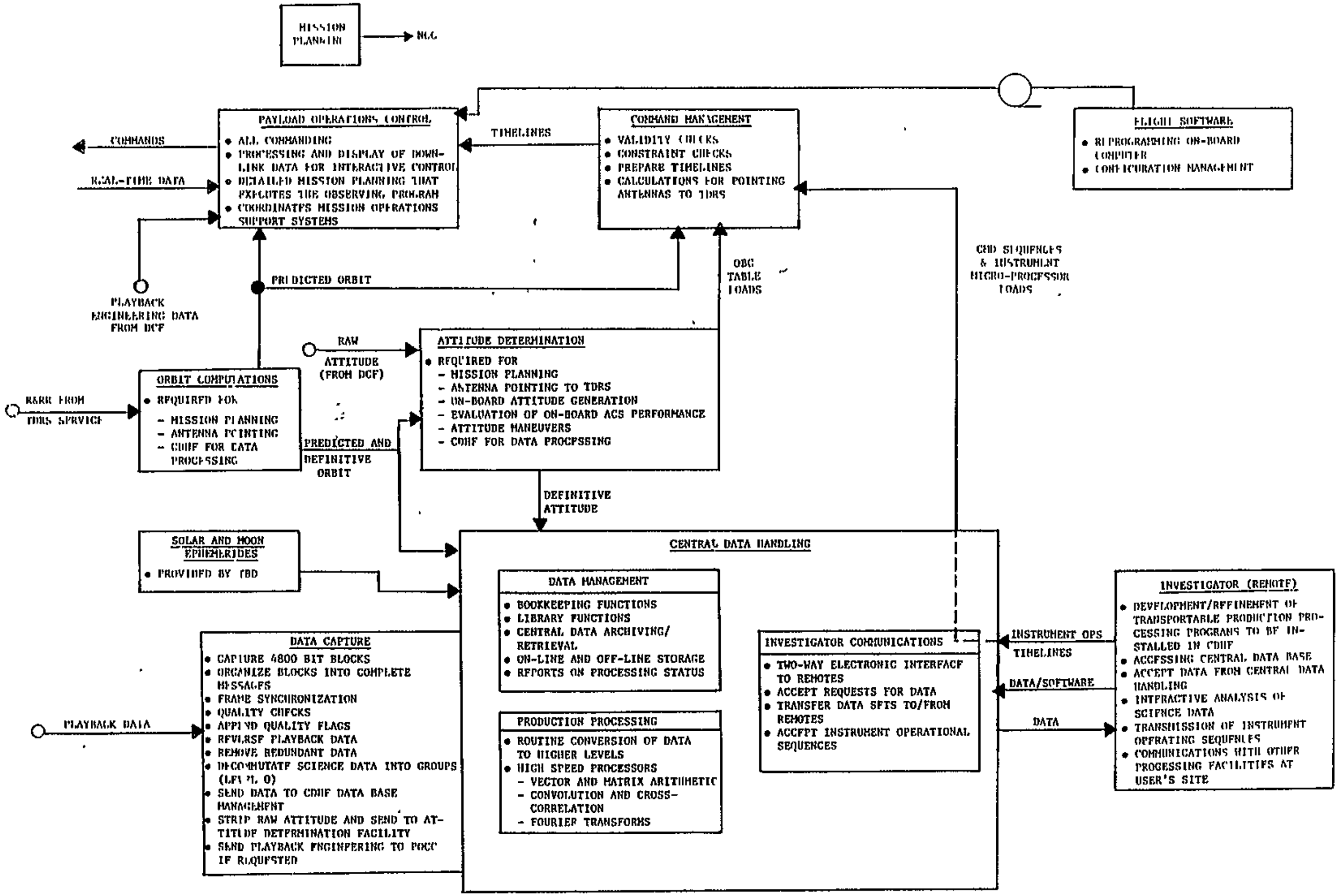



FIGURE 2.1-1 OPEN-UARS GROUND SYSTEM ELEMENTS

- Payload operations control
- Flight software
- Mission planning
- Communications

Additionally, the PI's will be provided interactive remote facilities suitable for analysis of the processed data.

The main functions performed by the ground system elements as well as the inter-relationships among them are shown in Figure 2.1-1.

# 3.0 DATA STORAGE AND PROCESSING ANALYSES

## 3.0 DATA STORAGE AND PROCESSING ANALYSES

In order to derive system concepts for processing and managing data within the CDHF, estimates for their data storage and data processing requirements are necessary. These are presented in Sections 3.1 and 3.2, respectively.

### 3.1 UARS and OPEN Storage Analysis

Information presented in the references (see Bibliography) allows for an analysis of the requirements for data processing and storage. It should be noted, however, that the information available for UARS is more complete.

Tables 3.1-1 through 3.1-3 present UARS data volumes for UARS production processing by data category. In addition, information regarding support data and PI data submissions from remote sites are included.

Table 3.1-4 presents the OPEN storage requirements. These requirements have been derived from information presented in the proposals for the OPEN instruments which have been selected.

### 3.2 UARS and OPEN Processing Estimates

In order to derive system concepts for the CDHF, not only must the data requirements be at hand but it is also necessary to focus upon the magnitude of the processing demands upon the CDHF. These are presented for UARS and OPEN in the following two sections, respectively.

TABLE 3.1-1 DAILY UARS INSTRUMENT DATA VOLUME

| INSTRUMENT | Avg. Data Rate (Kbps) | Daily Volume (MB) | | | | | Remarks |
|---|---|---|---|---|---|---|---|
| | | Level | | | | Total | |
| | | 0 | 1 | 2 | 3 | | |
| Winds and Temperatures (WINTERS) | 1.3 | 14 | 14.5 | 14.2 | 2.8 | 45.5 | |
| High Resolution Doppler Imager (HRDI) | 4.5 | 48.4 | 86 | 40 | 0.6 | 175.0 | |
| Cryogenic Limb Array Etalon Spectrometer (CLAES) | 1.1 | 12 | 32 | 10.9 | 0.18 | 45.08 | |
| Halogen Occultation Experiment (HALOE) | 1.1 | 15.8 | 12.5 | 2.7 | 0.04 | 31.04 | |
| Improved Stratospheric and Mesospheric Sounder (ISAMS) | 0.5 | 5.4 | 2.7 | 0.8 | 8.3 | 17.2 | |
| Microwave Limb Scanner (MLS) | 4.0 | 58.8 | 90 | 61 | 31 | 240.8 | |
| Particle Environment Monitor (PEM) | 2.7 | 28.5 | 109.5 | 11* | 5* | 154.0 | Levels 2 & 3 volumes are estimated from PI's requirements for graphics data. |
| Solar Ultraviolet Spectral Irradiance Monitor (SUSIM) | 1.0 | 10.7 | 0.5 | 0.23 | 0.01 | 11.44 | |
| Solar Stellar Irradiance Compar. Exper. (SOLSTICE) | 0.1 | 0.7 | 0.5 | 0.5 | 0.02 | 1.72 | |
| Solar Backscatter Ultra-violet Radiometer (SBUV)** | 0.32 | 2.4 | 2 | 2 | 0.4 | 6.8 | |
| (MAGNETOMETER) | 0.3 | 3.25 | 6.5 | - | - | 9.75 | Will be supplied by PEM and used in the PEM experiment only. |
| TOTAL FROM ALL INSTRUMENTS (MB) | | 199.95 | 356.7 | 133.33 | 48.35 | 738.33 | |

* 10:1 decrease in data volume from L1 --> L2 and 2:1 decrease from L2 --> L3 (estimated)
** Similar to instrument flown on advanced TIROS-N Series.

TABLE 3.1-2 UARS INSTRUMENT ORIENTED STORAGE
(Megabytes)

| | PRODUCTION DATA | | | | | OTHER DATA | | |
|---|---|---|---|---|---|---|---|---|
| INSTRUMENT | L0 (10 DAYS) | L1 (30 DAYS) | L2 (540 DAYS) | L3 (540 DAYS) | SUBTOTAL | DATA SUBMISSIONS FROM REMOTES (>L3) (540 DAYS) | MINIMAL SUPPORT DATA [4] | GRAND TOTALS (540 DAYS) |
| WINTERS | 140 | 435 | 7,668 | 1,512 | 9,755 | 0.0 | 22.0 [5] | 9,777 |
| HRDI | 484 | 2,580 | 21,600 | 324 | 24,988 | 32.4 [3] | 0.1 | 25,020 |
| CLAES | 120 | 960 | 486 | 97 | 1,663 | 9.7 [3] | 22.0 [5] | 1,695 |
| HALOE | 158 | 375 | 1,458 | 22 | 2,013 | 45.2 | 3.1 | 2,061 |
| ISAMS | 54 | 81 | 432 | 4,482 | 5,049 | 540.0 | 10.1 | 5,599 |
| MLS | 588 | 2,700 | 32,940 | 16,740 | 52,968 | 8,141.4 | 113.4 | 61,223 |
| PEM | 317 [1] | 3,480 [2] | 5,940 | 2,700 | 12,437 | 45.0 [3] | 22.0 [5] | 12,504 |
| SUSIM | 107 | 15 | 124 | 5 | 251 | 0.1 [3] | 4.8 | 256 |
| SOLSTICE | 7 | 15 | 270 | 11 | 303 | 5.9 | 0.3 | 309 |
| SBUV | 24 | 60 | 1,080 | 216 | 1,380 | 3.6 [3] | 22.0 [5] | 1,406 |
| Totals: | 1,999 | 10,701 | 71,998 | 26,109 | 110,807 | 8,823.3 | 219.8 | 119,850 |

Notes:

1. L0/day = PEM L0(28.5) + Magnetometer L0(3.25)

2. L1/day = PEM L1(109.5) + Magnetometer L1(6.5)

3. No estimate given; 10% of L3 assumed

4. Calibration processing coefficients, ground truth measurements, laboratory measurements, other correlative measurements

5. No estimate given; value assigned is average of 6 estimates that were given ($1.3175 \times 10^8$ / 6 = $22 \times 10^6$)

ORIGINAL PAGE IS OF POOR QUALITY

TABLE 3.1-3. UARS PRODUCTION DATA STORAGE REQUIREMENTS

| Production Data Type | Sequential Days of Data Concurrently On-Line | | | Daily Data Quantity (M Bytes) | Concurrent On-Line Storage Requirements (M Bytes) | | | Archive Storage Requirements For 540 Days (M Bytes) |
|---|---|---|---|---|---|---|---|---|
| | On Shared Disk | In MSS | TOTAL | | On Shared Disk | In MSS | TOTAL | |
| L-0 | - | 10 | 10 | 199.95 | - | 1,999 | 1,999 | 107,973 |
| L-1 | 10 | 20 | 30 | 356.7 | 3,567 | 7,134 | 10,701 | 192,618 |
| L-2 | 10 | 530 | 540 | 133.33 | 1,333 | 70,665 | 71,998 | 71,998 |
| L-3 | 10 | 530 | 540 | 48.35 | 483 | 25,625 | 26,108 | 26,109 |
| | | | TOTALS | 738.33 | 5,383 | 105,423 | 110,806 | 398,698 |

ORIGINAL PAGE IS
OF POOR QUALITY

# TABLE 3.1-4 OPEN DATA STORAGE REQUIREMENTS

| | PRINCIPAL INVESTIGATOR (PI) | | | | | Down-Link Data Rate Kbps | Duty Cycle % | Average Data Rate Kbps | Daily Data Volume MB | DAILY DATA VOLUME (MB) | | | | On-Line Storage MB[6] | Off-Line Storage Requirement MB[7] | Need for Array Processor |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Name | Institution/ Organization | ID | Experiment | Instrument Name | | | | | L0 | L1[4] | L2[5] | Total | | | |
| | Russell | UCLA | 81 | MAG. FIELDS | | 0.5 | 100 | 0.5 | 5.4 | 5.4 | 16.2 | 3.24 | 24.84 | 1,944 | 18,133.2 | |
| | Mozer | UCB | 85 | ELEC. FIELDS | | 2.5<br>5.0 | 90<br>10 | 2.75 | 29.7 | 29.7 | 89.1 | 17.82 | 136.62 | 10,692 | 99,732.6 | |
| | Shawhan | U of Iowa | 25 | PLASMA WAVE INSTRUMENT | | [3] | | 16.64 | 179.7 | 179.7 | 539.1 | 107.82 | 826.62 | 64,692 | 603,432.6 | |
| | Scudder | GSFC | 26 | HOT PLASMA | HYDRA | 4.4 | 100[2] | 4.4 | 47.5 | 47.5 | 142.5 | 28.5 | 218.5 | 17,100 | 159,505.0 | |
| | Shelley | LPARL | 24 | HOT PLASMA COMPOSITION | | 4.35 | 100[2] | 4.35 | 47.0 | 47.0 | 141.0 | 28.2 | 216.2 | 16,920 | 157,826.0 | |
| | Chappel[1] | MSFC | 31 | COLD PLASMA | TIDE | 2.5 | 100[2] | 2.5 | 27.0 | 27.0 | 81.0 | 16.2 | 124.2 | 9,720 | 90,666.0 | |
| P | Higbie[1] | LASL | 43 | ELEC. PARTICLE | CEPPAD | 3.2<br>9.6 | 90<br>10 | 3.84 | 41.47 | 41.47 | 124.4 | 24.88 | 190.75 | 14,928 | 139,247.5 | |
| P | Fritz[1] | NOAA/SEL | 53 | ELEC. PARTICLE COMPOSITION | CAMMICE | 1.152 | 100 | 1.152 | 12.45 | 12.45 | 37.35 | 7.47 | 57.27 | 4,482 | 41,807.1 | |
| L | Feldman | JHU | 90 | VISIBLE IMAGER | MAID | 16.2 | 80 | 12.96 | 140.0 | 140.0 | 420.0 | 84.0 | 644.0 | '50,400 | 470,120.0 | YES |
| | Torr | Utah | 92 | UV IMAGER | AIME | TBD | TBD | 20 | 216.0 | 216.0 | 648.0 | 129.6 | 993.6 | 77,760 | 725,328.0 | YES |
| | Imhof | LPARL | 64 | X-RAY IMAGER | PIXIE | 3.0 | 80 | 2.4 | 25.92 | 25.92 | 77.76 | 15.55 | 119.23 | 9,331 | 87,039.9 | |
| | | | | | | | LAB TOTAL | 71.49 | 772.1 | 772.1 | 2,316.3 | 463.3 | 3,551.66 | 277,956 | 2,592,711.8 | |
| | McPherron | UCLA | 66 | MAG. FIELDS | | 0.5 | 100 | 0.5 | 5.4 | 5.4 | 16.2 | 3.24 | 24.84 | 1,944 | 27,199.8 | |
| | Maynard | GSFC | 59 | ELEC. FIELDS | | 1.4<br>2.0 | 90<br>10 | 1.46 | 15.8 | 15.8 | 47.4 | 9.48 | 72.68 | 5,688 | 79,584.6 | |
| | McIlwain | UCSD | 74 | HOT PLASMA | EFIELDS | 2.0 | 100 | 2.0 | 21.6 | 21.6 | 64.8 | 12.96 | 99.36 | 7,776 | 108,799.2 | |
| | Scarf | TRW | 91 | PLASMA WAVE INSTRUMENT | | 2.138<br>6.158 | 90<br>10 | 2.54 | 27.44 | 27.44 | 82.32 | 16.46 | 126.22 | 9,878 | 138,210.9 | YES |
| | Parks | U of Wash. | 19 | HOT PLASMA | | 8.2 | 100[2] | 8.2 | 88.56 | 88.56 | 265.68 | 53.14 | 407.38 | 31,882 | 446,081.1 | |
| E | Burch | SRI | 45 | HOT PLASMA COMPOSITION | | 2.0<br>5.0 | 10<br>1 | 0.25 | 2.7 | 2.7 | 8.1 | 1.62 | 12.42 | 972 | 13,599.9 | |
| M | Chappell[1] | MCSF | 31 | COLD PLASMA | TIDE | 2.5 | 100[2] | 2.5 | 27.0 | 27.0 | 81.0 | 16.2 | 124.2 | 9,720 | 135,999.0 | |
| L | Higbie[1] | LASL | 43 | ELEC. PARTICLE | CEPPAD | 3.2<br>9.6 | 90<br>10 | 3.84 | 41.47 | 41.47 | 124.4 | 24.88 | 190.75 | 14,928 | 208,871.25 | |
| | Fritz[1] | NOAA/SEL | 53 | ELEC. PARTICLE COMPOSITION | CAMMICE | 1.464 | 100 | 1.464 | 15.8 | 15.8 | 47.4 | 9.48 | 72.68 | 5,688 | 79,584.6 | |
| | | | | | | | LAB TOTAL | 22.75 | 245.7 | 245.7 | 737.1 | 147.42 | 1,130.22 | 88,452 | 1,237,590.9 | |
| | Lepping | GSFC | 33 | MAG. FIELDS | | 0.65<br>3.7 | 90<br>10 | 0.955 | 10.31 | 10.31 | 30.93 | 6.18 | 47.42 | 3,711 | 51,924.9 | |
| | Mozer[1] | UCB | 84 | ELEC. FIELDS | | 0.6<br>1.2 | 90<br>10 | 0.66 | 7.1 | 7.1 | 21.3 | 4.26 | 32.66 | 2,556 | 35,762.7 | |
| G | Gurnett | U of Iowa | 46 | PLASMA WAVE INSTRUMENT | | 1.174<br>10.0 | 100<br>5 | 1.674 | 18.08 | 18.08 | 54.24 | 10.85 | 83.17 | 6,509 | 91,071.15 | YES |
| T | Frank | U of Iowa | 77 | HOT PLASMA | LEPEDEA | 3.5 | 100 | 3.5 | 37.8 | 37.8 | 113.4 | 22.68 | 173.88 | 13,608 | 190,398.6 | |
| L | Williams | NOAA/SEL | 63 | ELEC. PARTICLE COMPOSITION | EPIC | 2.7 | 100 | 2.7 | 29.16 | 29.16 | 87.48 | 17.49 | 134.13 | 10,497 | 146,872.35 | |
| | | | | | | | LAB TOTAL | 9.49 | 102.4 | 102.4 | 307.2 | 61.44 | 471.04 | 36,864 | 515,788.8 | |
| | Behannon | GSFC | 34 | MAG. FIELDS | | 0.65<br>3.7 | 90<br>10 | 0.955 | 10.31 | 10.31 | 30.93 | 6.18 | 47.42 | 3,711 | 51,924.9 | |
| I | Kaiser | GSFC | 35 | PLASMA WAVE INSTRUMENT | | 0.515<br>6.96 | 90<br>10 | 1.16 | 12.52 | 12.52 | 37.56 | 7.51 | 57.59 | 4,507 | 63,061.05 | |
| P | Ogilvie | GSFC | 50 | HOT PLASMA | | 0.40 | 100[2] | 0.40 | 4.32 | 4.32 | 12.96 | 2.59 | 19.87 | 1,555 | 21,757.65 | YES |
| | Gloeckler | U of MD | 86 | HOT PLASMA COMPOSITION | | 0.471 | 100[2] | 0.471 | 5.08 | 5.08 | 15.24 | 3.05 | 23.37 | 1,829 | 25,590.15 | |
| L | Lin | UCB | 13 | ELEC. PARTICLE | | 0.380 | 100[2] | 0.380 | 4.10 | 4.10 | 12.3 | 2.46 | 18.86 | 1,476 | 20,651.7 | |
| | McDonald | GSFC | 54 | COSMIC RAYS | EPACT | 0.25 | 100 | 0.25 | 2.7 | 2.7 | 8.1 | 1.62 | 12.42 | 972 | 13,599.9 | |
| | Teegarden | GSFC | 28 | GAMMA RAYS | | 0.193 | 100 | 0.193 | 2.08 | 2.08 | 6.24 | 1.25 | 9.57 | 749 | 10,479.15 | |
| | | | | | | | LAB TOTAL | 3.81 | 41.1 | 41.1 | 123.3 | 24.66 | 189.06 | 14,796 | 207,020.7 | |
| | | | | | | TOTALS FOR ALL FOUR LABS | | 107.54 | 1,161.3 | 1,161.3 | 3,483.9 | 698.78 | 5,341.98 | 418,068.0 | 4,553,112.2 | |

[1] Repeats
[2] Assumed

[3] 1.5 Kbps @ 100% duty cycle; 35.2 Kbps @ 100%; and 128 Kbps @ 5% (256 Kbps for 30 minutes; or 25.6 KHz for 4 hours)

[4] Increase of 3:1 from L0 → L1 (assumed)
[5] Decrease of 5:1 from L1 → L2 (assumed)
[6] 100 Days L1 + 100 Days L2
[7] 36 months each for EML, GTL, IPL; 24 months of FPL

### 3.2.1 UARS Processing Estimates

Based upon the data processing requirements contained in the actual questionnaire responses submitted by the PIs and the CSC study which summarizes and synthesizes these responses (References 6 and 7 of the UARS Bibliography), it is estimated that in order to process a day's data for the instruments selected (Appendix B) a total load of about 97,000 seconds of processing (excluding I/O) is required for computing machinery with an effective throughput of 0.5 Million Floating Point Operations (MFLOPS) per second. Thus,

$$\frac{0.5 \text{ MFLOPS}}{\text{sec}} \times 97{,}000 \text{ sec} = 48{,}000 \text{ MFLOPS}$$

are estimated for a day's production run. If these operations could be spread uniformly over an 8-hour period (one shift) then the effective throughput of the computing machinery would be:

$$48{,}500 \text{ MFLOPS} \times \frac{1}{8 \text{ hr}} \times \frac{1 \text{ hr}}{3600 \text{ sec}} = 1.68 \text{ MFLOPS/sec.}$$

In other words, CPU sizing for processing should be in the 2 MFLOPS/sec (effective throughput) range. Note that I/O and data management demands are not included.

### 3.2.2 OPEN Processing Estimates

Information for OPEN data processing which is comparable to the results in the CSC study has not yet been developed. However, gross estimates can be made for OPEN by extrapolating what is known about UARS together with analyzing the selected OPEN instrument proposals. When this is done, it is estimated that CPU sizing for OPEN data pro-

cessing is about in the 9 MFLOPS/sec range (effective throughput), excluding I/O and data management demands. The analysis for deriving this number is as follows:

For UARS the data being "transformed" during routine production processing are L-0, L-1 and L-2 which are transformed into L-1, L-2 and L-3, respectively. For OPEN, L-0 and L-1 data are transformed into L-1 and L-2, respectively. The daily quantity of data being transformed is as follows:

UARS: L-0 + L-1 + L-2 = 690 MB/day

OPEN: L-0 + L-1 = 4645 MB/day

See Tables 3.1-1 and 3.1-4, respectively.

For UARS, transforming the following instruments' L-1 data into L-2 accounts for about 91% of the total routine production processing demands: MLS, HALOE, ISAMS, CLAES, and WINTERS. Their associated quantity of L-1 data transformed daily is 151.7 Mbytes (see Table 3.1-1), which represents 151.7/690 = 22% of the total data which is transformed. These instruments were chosen as candidates for array processing. (Their processing demand estimates are reflected in the analysis presented in Section 4.3.)

For OPEN, five instruments are chosen as having similar processing demands as the UARS instruments indicated in the previous paragraph. These are the instruments of Feldman, Torr, Scarf, Gurnett, and Ogilvie. Their associated quantity of L-1 data transformed daily is 798 MB (see Table 3.1-4), which represents 798/4645 = 17% of the total data to be transformed. Note that this is comparable to the

analogous UARS percentage. The following assumptions are therefore made:

A1) Transforming the L-1 data of the preceding 5 OPEN instruments will account for about 90% of the OPEN routine processing demands.

A2) The processing demands for transforming the L-1 data of the preceding 5 OPEN instruments is about 798/152 = 5.25 times the processing demands for transforming the L-1 data of the 5 UARS instruments listed previously.

From Section 3.2.1, it was estimated that CPU sizing for UARS production processing is a minimum of 1.68 MFLOPS/sec (effective throughput). Thus, about 90% of 1.68 MFLOPS = 1.51 MFLOPS/sec are required to process the 151.7 MB of L-1 data of the selected UARS instruments. Therefore, from Assumption 2, 5.25 times 1.51 = 7.93 MFLOPS/sec would be required to process the analogous 798 MB of OPEN data. Adding 10% for the remaining production processing yields an estimate of 7.93 + 0.79 = 8.72 MFLOPS/sec to process a day's worth of OPEN data in an eight-hour period, excluding I/O and data management demands. The ratio of OPEN processing demands to UARS processing demands is 8.72:1.68 = 5.2:1.

# 4.0 CDHF SYSTEM CONCEPTS

## 4.0 CDHF SYSTEM CONCEPTS

This section presents two system design approaches for satisfying the requirements of either a UARS or an OPEN CDHF. Because the UARS CDHF is assumed to precede the OPEN CDHF, the overall approach has been to size a system and select components for a UARS CDHF, but in a manner that does not optimize the CDHF for UARS at the expense of OPEN. Indeed, the shortfall in system capability for OPEN support could be remedied by component upgrades rather than by a system re-design.

In what follows, a detailed analysis is made for UARS. System upgrades to accommodate OPEN are discussed in Section 4.4.

Based upon available information, the following major UARS functions have been identified:

- Data Ingest and L-0 Production
- L-0 to L-1 Production
- L-1 to L-2 Production
- L-2 to L-3 Production
- Data Services To/From Remotes
  - Browsing
  - Data File Transfers
- Remote Batch
  - Scheduling
  - Services
- Data Management

Based upon these functions, two functional concepts for a UARS CDHF have been formulated. The first concept presented is a CDHF featuring dual mainframe systems. The second concept presented is a CDHF confi-

guration featuring a single mainframe system. Neither concept depends upon unique hardware subsystems available from only a single vendor. The dual mainframe and single mainframe concepts are described in Sections 4.1 and 4.2, respectively, and two different hardware implementations of each concept are presented. Summary information regarding significant features and costs are presented in Tables 4.0-1 and 4.0-2.

## 4.1 Dual Mainframe Concept

In the dual mainframe concept the various CDHF functions are carried out by two autonomous software compatible mainframes which share a common database, and the CDHF functional workload is split between a Production Processor (PP) system and a Data Manager/Processor (DM/P) system as illustrated in Figure 4.1-1. As indicated in Figure 4.1-1, the extensive arithmetic and matrix manipulation services required to accomplish daily L-2 production and to provide remote batch services are provided by the PP and its associated array processing facilities, while the computationally less demanding L-0, L-1 and L-3 production services, as well as the (primarily) non-arithmetic data ingestion, data management and remote site interface services are provided by the DM/P.

The PP and DM/P would be sized to permit the processing of a day's volume of UARS data in one work shift, with capacity to spare. The PP would be sized in the 3 to 3.5 MIPS range, while the less powerful DM/P would operate in the range of 1 MIPS.

Since the dual mainframe concept features two independent software compatible mainframes sharing a common database, certain backup

TABLE 4.0-1

ALTERNATE IMPLEMENTATION FEATURE SUMMARY

| DUAL MAINFRAME IMPLEMENTATION | | SINGLE MAINFRAME IMPLEMENTATION | |
|---|---|---|---|
| IBM[1] | CDC[2] | IBM[3] | CDC[4] |
| * Production Processor<br>• IBM 3033-N-8<br>- 8 Mbyte Memory<br>- 3 MIPS | * Production Processor<br>• CDC Cyber 170 Series 700 170-730 Dual Processor<br>- 262K x 60 Bit Memory<br>- 3.5 MIPS | | |
| • 2 FPS AP-109L Array Processors<br>- 6 MFLOPS Average (ea)<br>- 256K Memory (ea) | • CDC Advanced Flexible Processor (For Array Processing)<br>- 200 Million Arithmetic Operations/Second (Avg) | * System Processor<br>• IBM 3081<br>- 16 Mbyte Memory<br>- 10.4 MIPS | * System Processor<br>• CDC Cyber 170 Series 700 170-760 Processor<br>- 262K x 60 Bit Memory<br>- 11 MIPS<br>• Extended Memory<br>- 1 Million 60 Bit Words |
| * Data Manager/Production Processor<br>• IBM 4341-L02<br>- 8 Mbyte Memory<br>- 0.75 MIPS | * Data Manager/Production Processor<br>• CDC Cyber 170 Series 700 170-720 Processor<br>- 262K x 60 Bit Memory<br>- 1.2 MIPS<br>* Shared Extended Memory<br>• 1 Million 60 Bit Words | | |
| * On-Line Mass Storage[5] (Shared, not disk)<br>• 2 IBM 3851-A04<br>- 472 Gbytes Total | * On-Line Mass Storage[6] (Shared, not disk)<br>• 2 MASSTOR 860<br>- 440 Gbytes Total | * On-Line Mass Storage[5] (not disk)<br>• 2 IBM 3851-A04<br>- 472 Gbytes Total | * On-Line Mass Storage[6] (not disk)<br>• 2 MASSTOR M860<br>- 440 Gbytes Total |

Notes:

[1] Partial listing; complete listing in Table 4.1-3.

[2] Partial listing; complete listing in Table 4.1-6.

[3] Partial listing; complete listing in Table 4.2-3.

[4] Partial listing; complete listing in Table 4.2-5.

[5] System limit; may not be expanded.

[6] Expandable.

TABLE 4.0-2

SYSTEM COST SUMMARY ESTIMATES

| FACILITY | DUAL MAINFRAME IMPLEMENTATION | | SINGLE MAINFRAME IMPLEMENTATION | |
|---|---|---|---|---|
| | IBM[1] | CDC[2] | IBM[3] | CDC[4] |
| Central Data Handling Facility (CDHF) | $8,463,662 | $7,405,105 | $9,539,135 | $8,236,388 |
| Software Compatible Remote Sites[5] (1 site/19 sites) | $208,890/ $3,968,910 | $747,837/ $14,208,903 | $208,890/ $3,968,910 | $747,837/ $14,208,903 |
| Combined Cost of CDHF and 19 Remote Sites | $12,432,572 | $21,614,008 | $13,508,045 | $22,445,291 |

<u>Notes</u>:

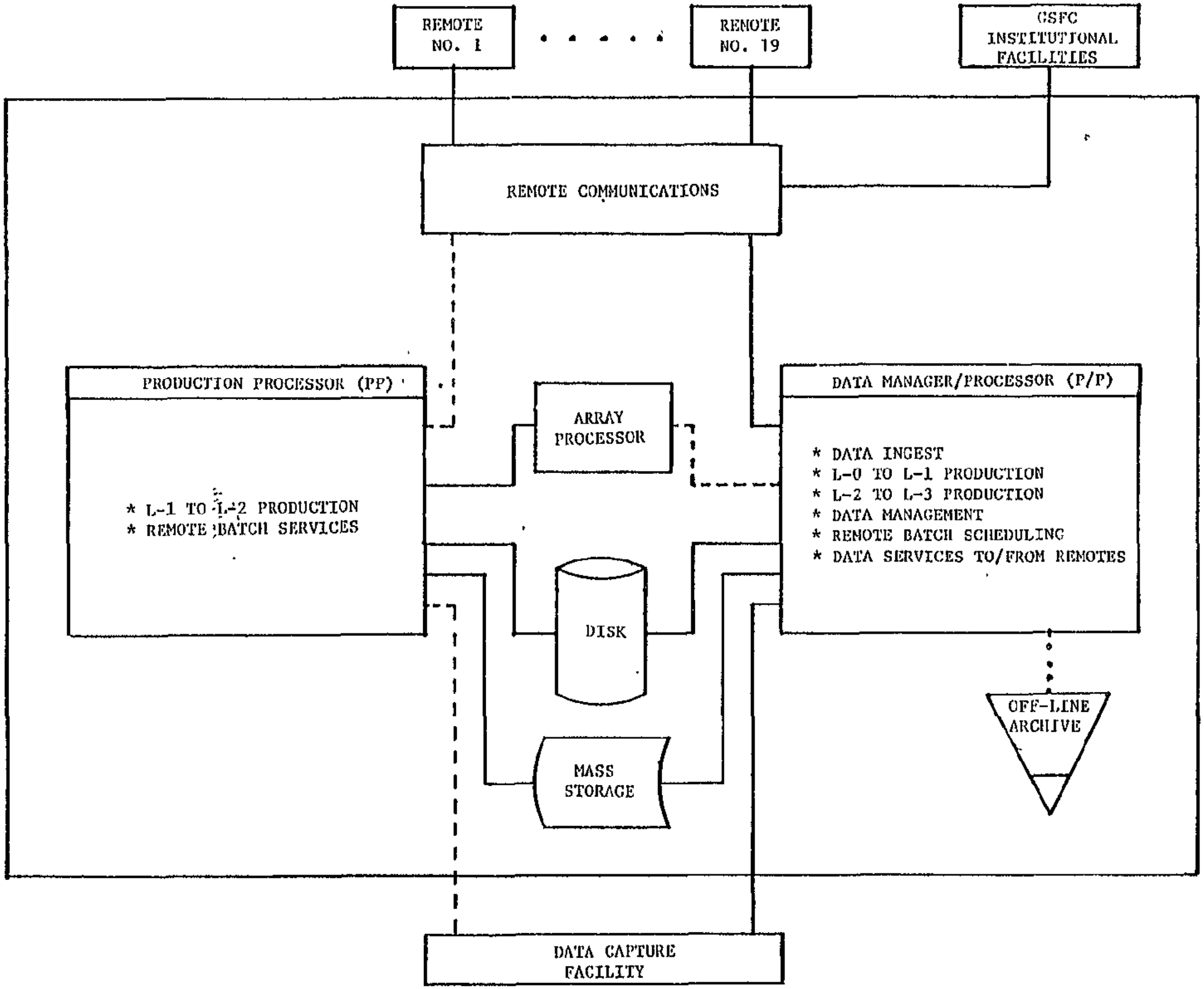
[1] Detailed cost estimates presented in Tables 4.1-3 and 4.1-4.

[2] Detailed cost estimates presented in Tables 4.1-6 and 4.1-7.

[3] Detailed cost estimates presented in Tables 4.2-3 and 4.1-4.

[4] Detailed cost estimates presented in Tables 4.2-5 and 4.1-7.

[5] CDC remote facilities have extensive computational capabilities appropriate for OPEN. IBM remote facilities, while less powerful, are more appropriate for UARS. The CDC remote facilities represent the low end of the software compatible Cyber 170 Series 700 equipment line.




LEGEND:
------ DENOTES BACKUP (REDUNDANT) PATH

FIGURE 4.1-1 DUAL MAINFRAME FUNCTIONAL CONCEPT

capabilities are inherent in this approach which are not present in a single mainframe approach. In the event of PP outage, the DM/P and array processing facilities may be used to carry on UARS production at a reduced rate of approximately 50% (2 work shifts, with little or no margin). In the event of DM/P outage, the PP can assume the responsibilities of the DM/P and complete all daily processing tasks within 2 work shifts.

Two possible hardware implementations of the dual mainframe concept have been prepared. The first implementation features dual IBM mainframes (Section 4.1.1), while the second implementation features dual CDC mainframes (Section 4.1.2). Table 4.1-1 summarizes these two implementations.

### 4.1.1 An IBM Hardware Implementation

This implementation of the dual mainframe concept is configured using hardware produced by the IBM Corporation and Floating Point System (FPS) Corporation. See Table 4.1-2. Figure 4.1-2 presents the general structure of this implementation. A more detailed illustration of this implementation is presented in Figure 4.1-3.

Cost summary information for an IBM dual mainframe CDHF and the corresponding remote (PI) facilities is presented in Tables 4.1-3 and 4.1-4, respectively.

### 4.1.2 A CDC Hardware Implementation

This implementation of the dual mainframe concept is configured using hardware produced by Control Data Corporation (CDC) and Masstor Systems Corporation. See Table 4.1-5. Figure 4.1-4 presents the general structure of this implementation.

TABLE 4.1-1

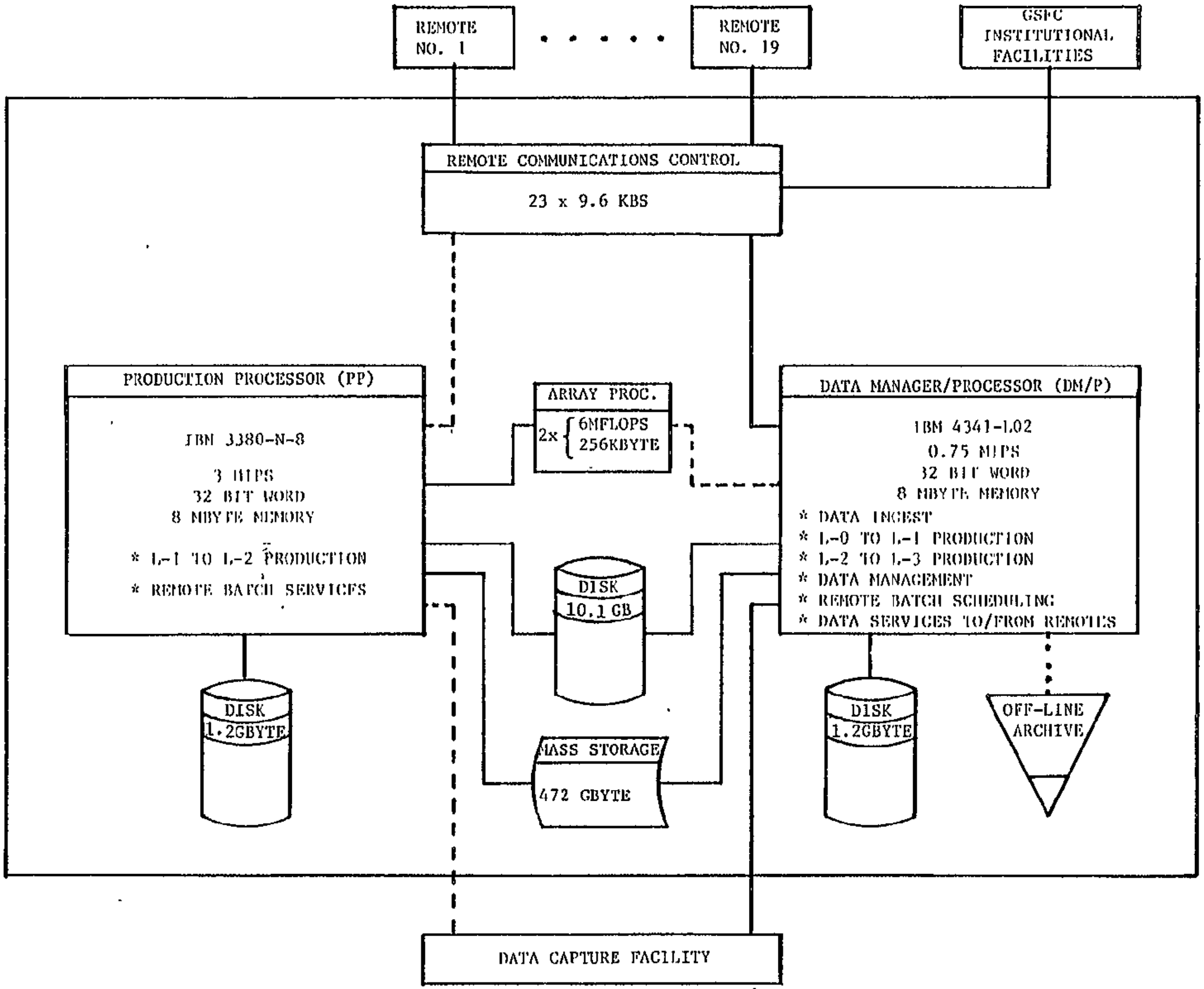
DUAL MAINFRAME IMPLEMENTATIONS

| Feature | IBM | CDC |
|---|---|---|
| Production Processor (PP) | IBM 3033-N-8 | CDC Cyber 170 170-730 Dual Processor |
| Data Manager/Production Processor (DM/P) | IBM 4341-L02 | CDC Cyber 170 170-720 Processor |
| Array Processor(s) | 2 Floating Point Systems AP-190L Array Processors | CDC Advanced Flexible Processor |
| Inter-Processor Data Exchange | Shared Disk | Shared Disk |
| Dedicated Separate System Disks | Yes | No |
| Mass Storage Facilities (other than disk) | IBM 3850 (472 Billion Bytes) | MASSTOR M-860 (440 Billion Bytes) |
| PP and DM/P Software Compatibility | Yes | Yes |

TABLE 4.1-2

AN IBM DUAL MAINFRAME SUBSYSTEM SUMMARY

| Subsystem | Manufacturer | Comments |
|---|---|---|
| Mainframe Processors | IBM | PP: Model 3380-N-8<br>DM/P: Model 4341-L02 |
| Tape | IBM | 2 Drives/Mainframe |
| Disk | IBM | 1.2 Gbytes (not shared)/ Mainframe;<br>10.1 Gbytes shared between mainframes (approximately 50% used for recent production data) |
| Communications | IBM | DM/P Subsystem |
| CDHF Terminals | IBM | 5 CRT Terminals and<br>2 Printers/Mainframe |
| Array Processor | Floating Point Systems | Off-the-shelf interface software readily available |
| On-Line Mass Storage | IBM | Shared; cannot be expanded beyond 472 Gbyte |




LEGEND:

------ DENOTES BACKUP (REDUNDANT) PATH

FIGURE 4.1-2 AN IBM DUAL MAINFRAME STRUCTURAL SUMMARY

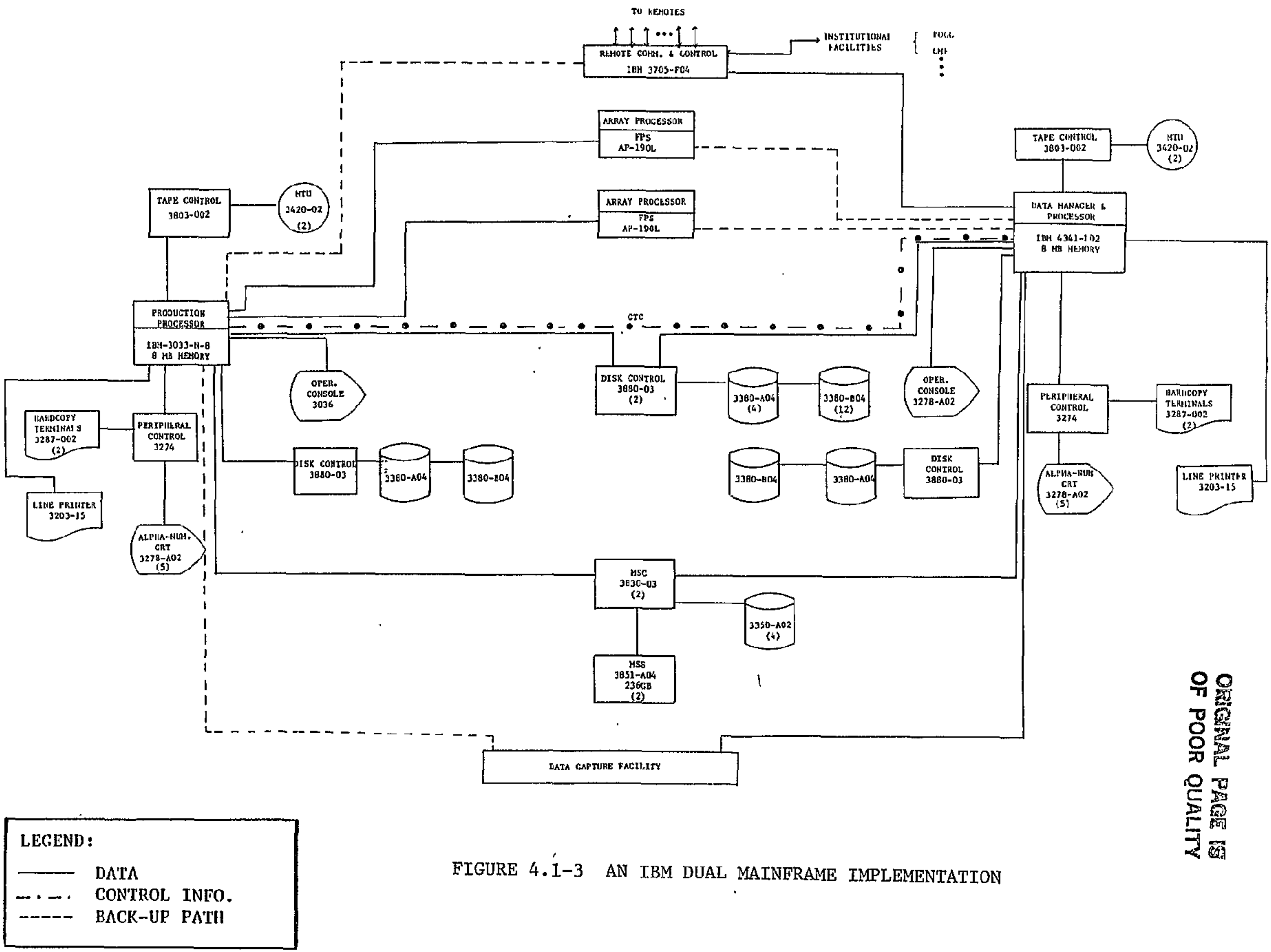



FIGURE 4.1-3 AN IBM DUAL MAINFRAME IMPLEMENTATION

ORIGINAL PAGE IS OF POOR QUALITY

TABLE 4.1-3

COSTS OF AN IBM DUAL MAINFRAME CDHF IMPLEMENTATION

| ITEM | GSA PURCHASE PRICE ($) |
|---|---|
| PRODUCTION PROCESSOR (PP) | |
| • IBM-3033-N-8 | 1,561,000 |
| - 3 MIPS<br>- 8 Mbyte Memory<br>- 12 Channels<br>- Power Unit<br>- Operator Console | |
| • Array Processors | 352,000 |
| - 2 FPS AP-190L<br>- 6 MFLOPS Average (each)<br>- 256 K Memory (each) | |
| • PP Disks | 268,360 |
| - 1.2 Billion Bytes<br>- 1 3880-003 Controller<br>- 1 3880-A04 Disk<br>- 1 3880-B04 Disk | |
| • PP Tapes | 85,175 |
| - 1 3803-002 Controller<br>- 2 3420-006 Tape Units (125 ips, 1600/6250 bpi) | |
| • Terminals | 40,524 |
| - 1 3274 Control Unit<br>- 5 3278 KB/CRT's<br>- 2 3287 Printers (friction feed) | |
| • Line Printer | 41,250 |
| - 1 3203-005 (1200 lpm, train cartridge) | |
| PP TOTAL COST | 2,348,309 |

TABLE 4.1-3 (Continued)

COSTS OF AN IBM DUAL MAINFRAME CDHF IMPLEMENTATION

| ITEM | GSA PURCHASE PRICE ($) |
|---|---|
| DATA MANAGER AND PROCESSOR (DM/P) | |
| • IBM-4341-L02 | 497,000 |
| - 8 Mbyte Memory<br>- 0.75 MIPS | |
| • DM/P Disks | 268,360 |
| - 1.2 Billion Bytes<br>- 1 3880-003 Controller<br>- 1 3380-A04 Disk<br>- 1 3380-B04 Disk | |
| • DM/P Tapes | 85,175 |
| - 1 3803-002 Controller<br>- 2 3420-006 Tape Units (125 ips, 1600/6250 bpi) | |
| • Terminals | 40,524 |
| - 1 3274 Control Unit<br>- 5 3278 KB/CRT's<br>- 2 3287 Printers (friction feed) | |
| • Line Printer | 41,250 |
| - 1 3203-005 (1200 lpm, train cartridge) | |
| • Communications | 86,890 |
| - 1 3705-F04 (24 Bi-sync lines @ 9.6 Kbps) | |
| DP/M TOTAL COST | 1,019,199 |

TABLE 4.1-3 (Concluded)

COSTS OF AN IBM DUAL MAINFRAME CDHF IMPLEMENTATION

| ITEM | GSA PURCHASE PRICE ($) |
|---|---|
| SHARED SUBSYSTEMS | |
| • Disk | 1,721,440 |
| - 10.144 Billion Bytes (Total)<br>- 2 3880-03 Controllers<br>- 4 3380-A04 Disks<br>- 12 3380-B04 Disks | |
| • Mass Storage System | 3,374,714 |
| - 472 Billion Bytes (Total)<br>2 3851-A04 Mass Storage Facilities (MSF)<br>236 Billion Bytes (each MSF)<br>4 Data Recording Controls (each MSF)<br>8 Data Recording Devices (each MSF)<br>4720 Cartridges (each MSF) @ $35 each<br>- 2 3830-003 Storage Control Units<br>- 2.536 Billion Bytes Staging Disk<br>(2 3350-A02, 2 3350-B02) | |
| SHARED SUBSYSTEM COST | 5,096,154 |
| CDHF TOTAL COST | 8,463,662 |

TABLE 4.1-4

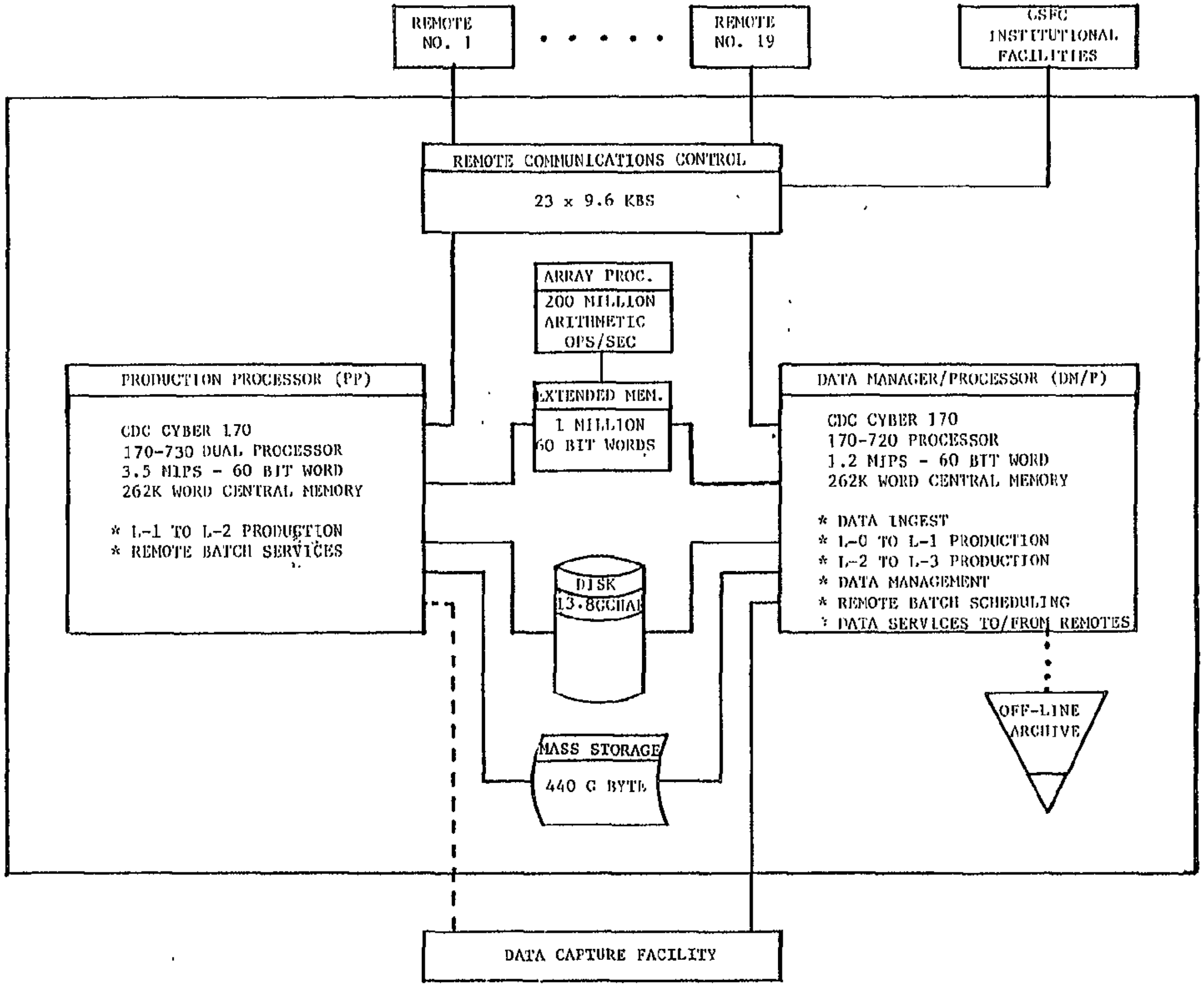
COSTS FOR IBM COMPATIBLE REMOTE FACILITIES

| ITEM | GSA PURCHASE PRICE ($) |
|---|---|
| • 1 IBM-4331-I01<br>- 512 KB Memory<br>- 0.5 MIPS<br>- 600 MB Disk Storage<br>- Communication Adapters | 118,750 |
| • 1 Graphics Terminal (Tektronics 4012) with Hardcopy Device (Tektronics 4631) | 30,000 |
| • 2 Magtape Controller and Tape Units<br>- 800/1600 bpi | 40,000 |
| • 3 Consoles<br>- 1 OP. Console<br>- 2 Alphanumeric Terminals | 7,000 |
| • 1 400 lpm Printer (IBM-3289) | 13,140 |
| TOTAL COST for 1 System | 208,890 |
| TOTAL COST for 19 Systems | 3,968,910 |

TABLE 4.1-5

A CDC DUAL MAINFRAME SUBSYSTEM SUMMARY

| Subsystem | Manufacturer | Comments |
|---|---|---|
| Mainframe Processors | CDC | PP: Series 700 170-730 Dual Processors DM/P: Series 700 170-720. |
| Tape | CDC | 4 Drives (shared) |
| Disk | CDC | 13.8 Billion 6 bit characters (shared); approximately 50% used for recent production data |
| Communications | CDC | Shared |
| CDHF Terminals | CDC | 10 Shared CRT Terminals and 4 Shared Printers |
| Array Processor | CDC | Advanced Flexible Processor (AFP) |
| On-Line Mass Storage | MASSTOR | Shared; M-860 systems marketed and supported by CDC; expandable |




LEGEND:
------ DENOTES BACKUP (REDUNDANT) PATH

FIGURE 4.1-4 A CDC DUAL MAINFRAME STRUCTURAL SUMMARY

A more detailed illustration indicating the extensive dual access features of this implementation is presented in Figure 4.1-5.

Cost summary information for a CDC dual mainframe CDHF and the corresponding remote (PI) facilities is presented in Tables 4.1-6 and 4.1-7, respectively.

## 4.2 Single Mainframe Concept

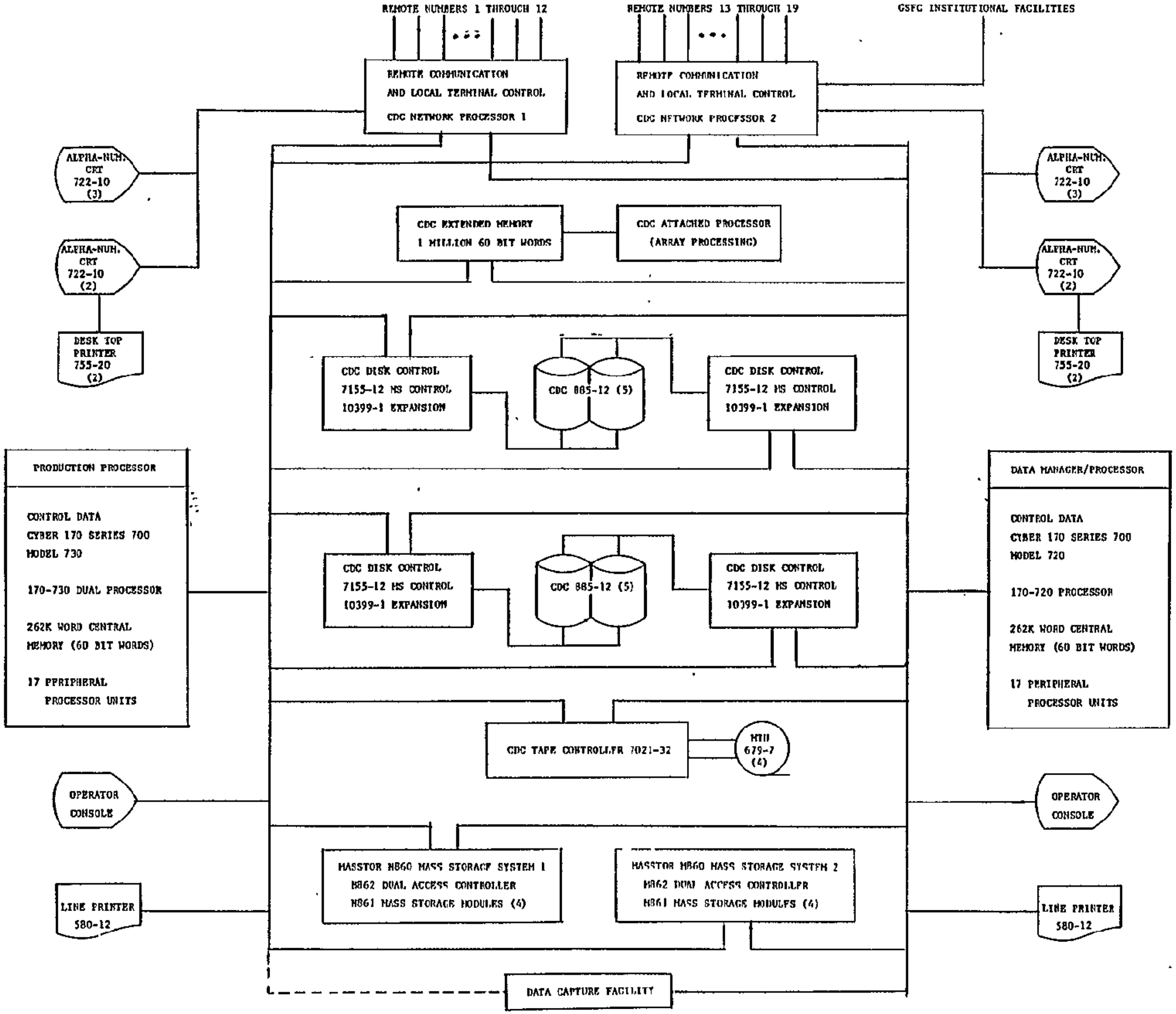
The single mainframe concept accomplishes all of the CDHF functions using a single large mainframe. This concept is illustrated in Figure 4.2-1.

As was the case with the dual mainframe concept, the single mainframe is sized to permit the processing of a day's volume of UARS data in one work shift. In contrast to the dual mainframe concept, however, the single mainframe concept does not include the capability to operate at reduced levels in the event of mainframe failure since there is no mainframe redundancy.

Sections 4.2.1 and 4.2.2 present hardware implementations of the single mainframe concept. A possible IBM implementation is described in Section 4.2.1, while Section 4.2.2 presents possible CDC implementation. Table 4.2-1 summarizes several features of these implementations.

### 4.2.1 An IBM Hardware Implementation

Table 4.2-2 summarizes several features of this all IBM system, while Figure 4.2-2 presents the general structure of this implementation. Figure 4.2-3 provides a more detailed illustration showing peripheral/channel relationships and device/controller cross-strapping for this implementation.




LEGEND:

------ DENOTES BACKUP (REDUNDANT) PATH

FIGURE 4.1-5 A CDC DUAL MAINFRAME IMPLEMENTATION

TABLE 4.1-6

COSTS OF A CDC DUAL MAINFRAME CDHF IMPLEMENTATION

| ITEM | GSA PURCHASE PRICE ($) |
|---|---|
| PRODUCTION PROCESSOR (PP) | |
| • CDC Cyber 170 Series 700 170-730 Dual Processor | 1,236,910 |
| - 3.5 MIPS | |
| - 60 Bit Word | |
| - 262K Word Central Memory | |
| - 24 Channels | |
| - 17 Peripheral Processors | |
| - Extended Memory Interface | |
| - Power Unit | |
| - Operator Console | |
| • Array Processor | 400,000 |
| - CDC Advanced Flexible Processor | |
| - 200 Million Arithmetic Operations/Seconds (AVG) | |
| • Terminals | 13,490 |
| - 5 Alphanumeric CRT's (CDC 722-10) | |
| - 2 Desktop Printers (CDC 755-20) | |
| • Line Printer | 60,187 |
| - 1200 lpm (CDC 580-12) | |
| - Printer Train Cartridge (CDC 596-6) | |
| PP TOTAL COST | 1,710,587 |

TABLE 4.1-6 (Continued)

COSTS OF A CDC DUAL MAINFRAME CDHF IMPLEMENTATION

| ITEM | GSA PURCHASE PRICE ($) |
|---|---|
| DATA MANAGER AND PROCESSOR (DM/P) | |
| • CDC Cyber Series 700 170-720 Processor | 801,535 |
| - 1.2 MIPS | |
| - 60 Bit Word | |
| - 262K Word Central Memory | |
| - 24 Data Channels | |
| - 17 Peripheral Processors | |
| - Extended Memory Interface | |
| - Power Unit | |
| - Operator Console | |
| • Terminals | 13,490 |
| - 5 Alphanumeric CRT's (CDC 722-10) | |
| - 2 Desk Top Printers (CDC 755-20) | |
| • Line Printer | 60,187 |
| - 1200 lpm (CDC 580-12) | |
| - Printer Train Cartridge (CDC 596-6) | |
| DM/P TOTAL COST | 875,212 |

TABLE 4.1-6 (Continued)

COSTS OF A CDC DUAL MAINFRAME CDHF IMPLEMENTATION

| ITEM | GSA PURCHASE PRICE ($) |
| --- | --- |
| **SHARED SUBSYSTEMS** | |
| • Extended Memory<br>- 1 Million 60 Bit Words<br>- 3 High Speed Ports<br>- 10-20 Million Words/Second | 663,200 |
| • Disk (CDC)<br>- 13.84 billion 6-bit characters<br>- 10 885-12 Disk Storage Units (dual spindle, two-controller)<br>- 4 7155-12 Two Channel Controllers<br>- 4 7155-885 Four Drive Expansion | 881,360 |
| • Tape (CDC)<br>- 4 679-7 Tape Transports (200 ips, 1600/6250 bpi)<br>- 1 7021-32 Dual Access Controller | 224,040 |
| • Mass Storage System (MASSTOR)<br>- 440 billion bytes<br>- 8 M861 Mass Storage Modules (16 read/write stations)<br>- 2 M862 Dual Access Storage Controllers<br>- 4 Channel Couplers (CDC 65206-X) | 2,911,800 |
| SHARED SUBSYSTEM TOTAL COST | 4,680,400 |

TABLE 4.1-6 (Concluded)

COSTS OF A CDC DUAL MAINFRAME CDHF IMPLEMENTATION

| ITEM | GSA PURCHASE PRICE ($) |
|---|---|
| COMMUNICATIONS | |
| • 2 Dual Access Subsystems | 138,906 |
| – 2 CDC 2551-2 Network Processing Units (NPU)<br>96K 16-bit words/NPU<br>2 558-3 Couplers/NPU<br>6 Sync. Comm Line Adapters (CLA) per NPU (2 Remote Lines/CLA)<br>3 Async. CLA's per NPU (2 Local CRT's/CLA) | |
| – Remote Site Interface<br>22 Bi-sync Lines (11/NPU)<br>9600 bps | |
| – PP CRT Terminal Interface (Hardwired)<br>5 Asynchronous Lines<br>9600 bps | |
| – DM/P CRT Terminal Interface (Hardwired)<br>5 Asynchronous Lines<br>9600 bps | |
| CDHF TOTAL COST | 7,405,105 |

TABLE 4.1-7

COSTS FOR CDC COMPATIBLE REMOTE FACILITIES

| ITEM | GSA PURCHASE PRICE ($) |
|---|---|
| • 1 CDC Cyber 170 Series 700 170-720 Processor<br>- 1.2 MIPS<br>- 60 Bit Word<br>- 98K Word Central Memory<br>- 10 Peripheral Processors<br>- Power Unit<br>- Operator Console | 479,545 |
| • Disk Subsystem<br>- 1 7155-11 Disk Controller<br>- 1 885-11 Dual Spindle Disk Storage Unit (1.384 billion characters) | 99,890 |
| • Tape Subsystem<br>- 1 7021-31 Tape Controller<br>- 2 679-2 Tape Transports (800/1600 bpi, 100 ips) | 81,720 |
| • Line Printer<br>- 1 1827-60 (600 lpm) | 17,000 |
| • 2 Terminals<br>- 2 Alphanumeric CRT's (CDC 722-10)<br>- Installation Charge | 3,000 |
| • 1 Graphics Terminal (Tektronics 4012) with Hardcopy Device (Tektronics 4631) | 30,000 |

TABLE 4.1-7 (Concluded)

COSTS FOR CDC COMPATIBLE REMOTE FACILITIES

| ITEM | GSA PURCHASE PRICE ($) |
|---|---|
| • Communications and Terminal Control<br>- CDHF Interface<br>- Local Alphanumeric (2) CRT Interface<br>- Local Graphics Terminal Interface<br>- 1 CDC 2551-1 Network Processing Unit (32K 16-bit word memory)<br>- 1 CDC 2580-4 Autostart Module-Cassette<br>- 1 Synchronous Comm. Line Adapter (2 Lines) (CDHF Interface, Graphics Terminal Interface)<br>- 1 Asynchronous Comm. Line Adapter (2 Lines) (Local Alphanumeric CRT Interface) | 36,682 |
| TOTAL COST for 1 System | 747,837 |
| TOTAL COST for 19 Systems | 14,208,903 |

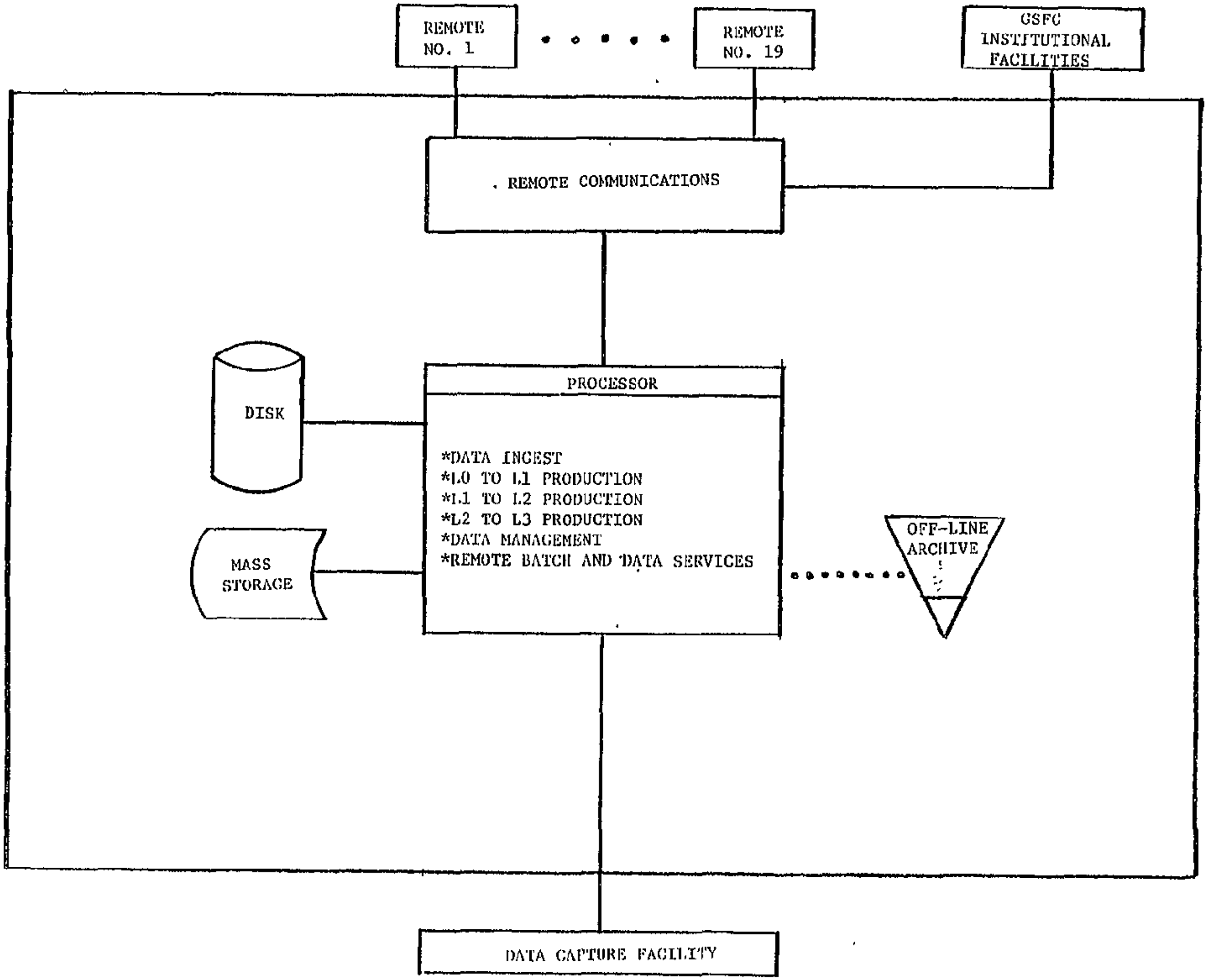



FIGURE 4.2-1 SINGLE MAINFRAME FUNCTIONAL CONCEPT

TABLE 4.2-1

SINGLE MAINFRAME IMPLEMENTATIONS

| Feature | IBM | CDC |
|---|---|---|
| System Processor | IBM 3081 | CDC Cyber 170 170-730 Processor |
| Array Processor | None | None |
| Mass Storage (other than disk) | IBM 3850 (472 Billion Bytes) | MASSTOR M-860 (440 Billion Bytes) |

TABLE 4.2-2

AN IBM SINGLE MAINFRAME SUBSYSTEM SUMMARY

| Subsystem | Manufacturer | Comments |
|---|---|---|
| Processor | IBM | Model 3081 |
| Tape | IBM | 4 Drives |
| Disk | IBM | 10.1 Gbytes; approximately 50% used for recent production data. |
| Communications | IBM | |
| CDHF Terminals | IBM | 10 CRT Terminals and 4 Printers |
| Array Processor | | None |
| On-Line Mass Storage | IBM | Cannot be expanded beyond 472 Gbyte |

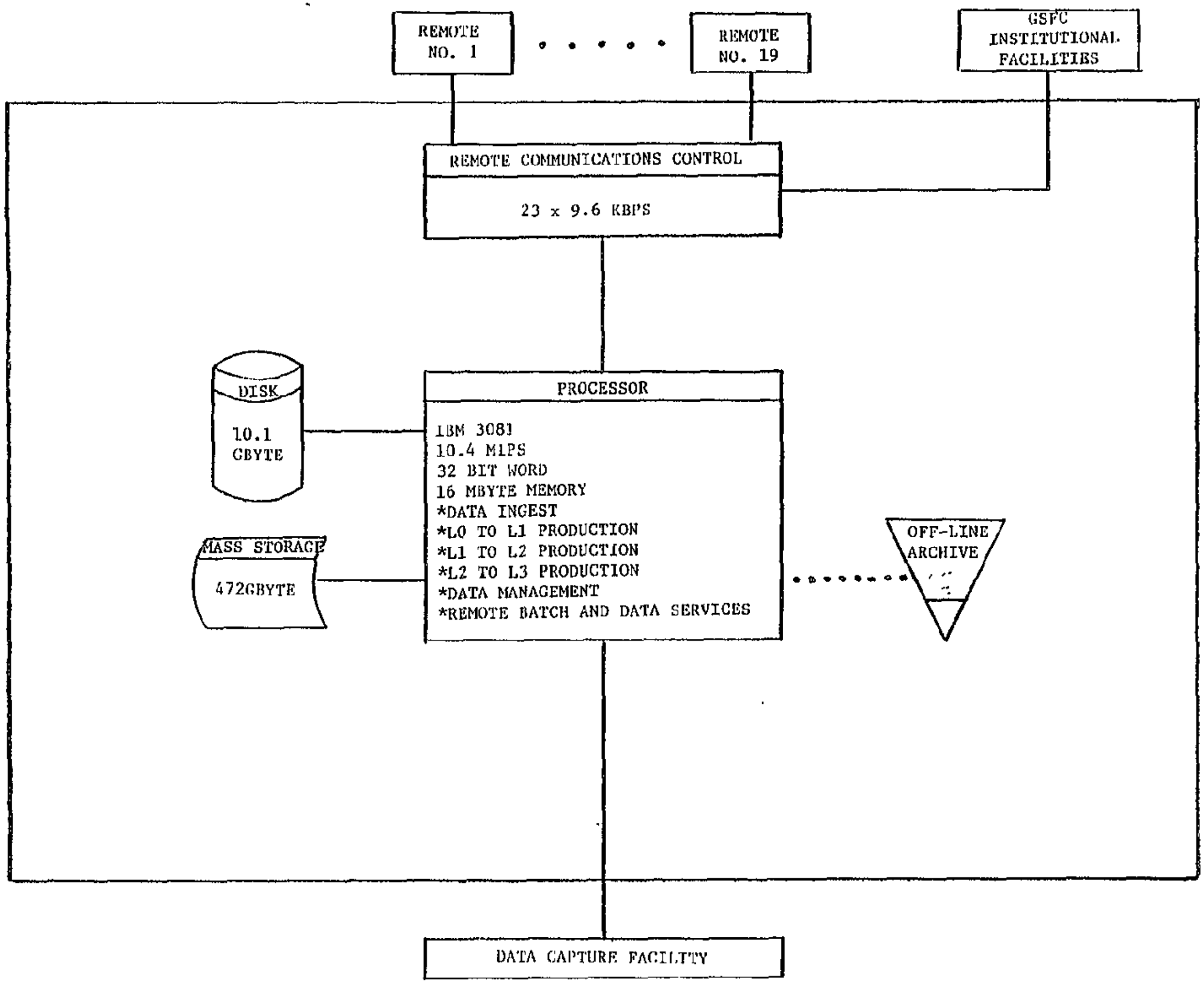



FIGURE 4.2-2 AN IBM SINGLE MAINFRAME STRUCTURAL SUMMARY

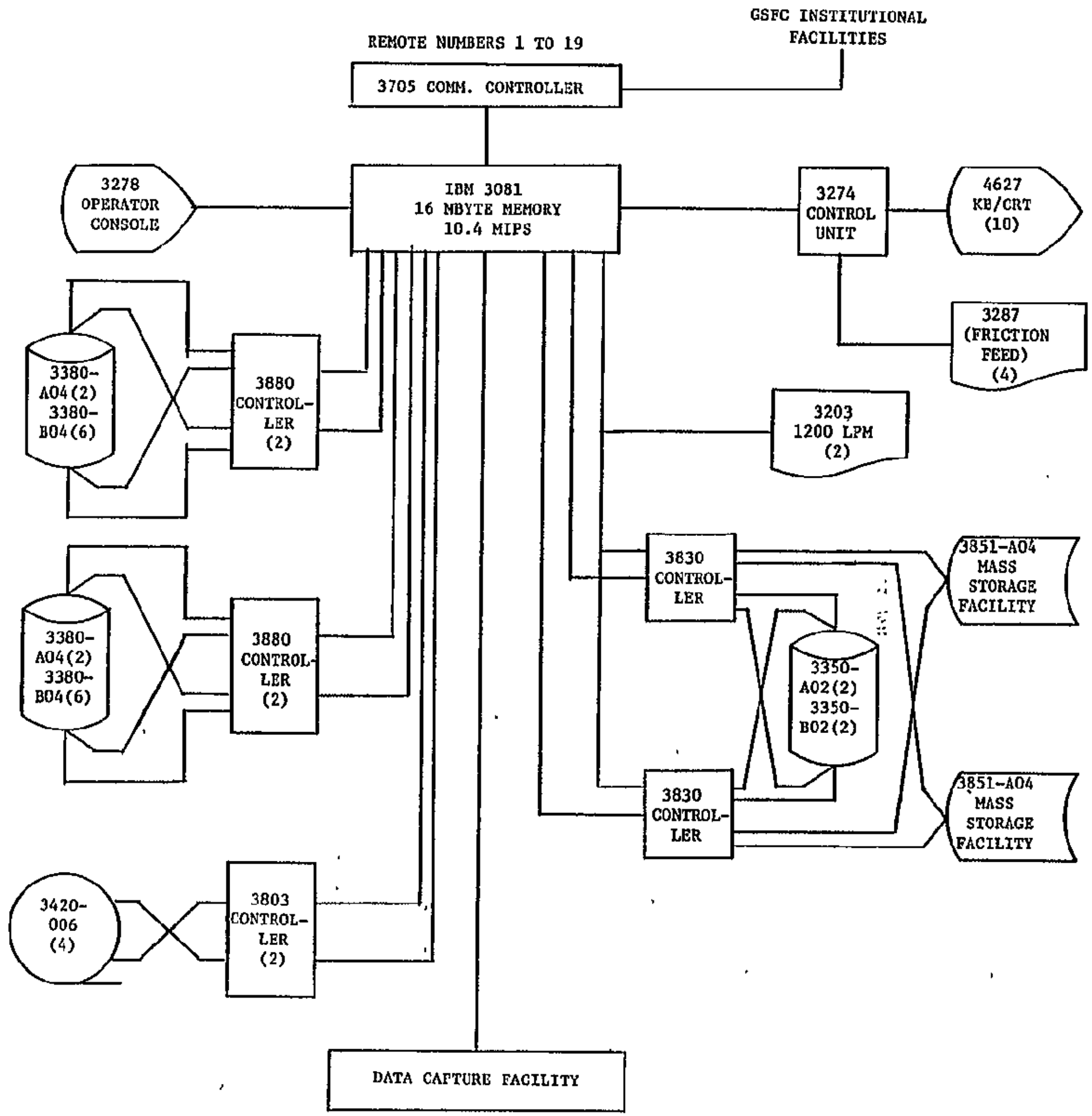



FIGURE 4.2-3 AN IBM SINGLE MAINFRAME IMPLEMENTATION

Cost summary information for the IBM single mainframe implementation is presented in Table 4.2-3. Costs for software compatible remote facilities are the same as those presented for the IBM dual mainframe system (see Table 4.1-4).

### 4.2.2 A CDC Hardware Implementation

Major elements of this CDC implementation are presented in Table 4.2-4. The major subsystems are illustrated in Figure 4.2-4. A detailed illustration of a possible CDC single mainframe CDHF implementation is presented in Figure 4.2-5.

Cost summary information for the CDC single mainframe CDHF is presented in Table 4.2-2. Costs for corresponding remote (PI) facilities would be as presented in Table 4.2-5.

Another option for a CDC single mainframe implementation would be to replace the 170-740 mainframe with a less powerful 170-740 processor and aray processing facilities (the CDC Advanced Flexible Processor). In contrast to the 170-760 processor, which operates in the range of 10-12 MIPS, the 170-740 operates in the range of 4.5 to 5 MIPS; the Advanced Flexible Processor (AFP) operates in the range of 200 million arithmetic operations per second.

The total cost of the 170-740/AFP system would be $6,931,788 in contrast to the $8,236,388 cost of a 170-760 total CDHF system. While this cost difference of $1,304,600 is not trivial, the introduction of array processing facilities external to the mainframe processor could increase software development costs at the CDHF and, especially, at the remote sites (which are not scheduled to have special array processing facilities).

TABLE 4.2-3

COSTS OF A IBM SINGLE MAINFRAME CDHF IMPLEMENTATION

| ITEM | GSA PURCHASE PRICE ($) |
|---|---|
| • IBM 3081 Processor | 4,040,874 |
| - 10.4 MIPS<br>- 16 MByte Memory<br>- 16 Channels<br>- Power Unit<br>- Coolant Distribution Unit<br>- Operator Console | |
| • Disk | 1,721,440 |
| - 10.144 Billion Bytes (Total)<br>- 2 Disk Subsystems (DSS)<br>- Each DSS<br>2 3880-003 Controllers<br>2 3380-A04 Disks<br>6 3380-B04 Disks | |
| • Tape | 170,350 |
| - 4 3420-006 Tape Units<br>(125 ips, 1600/6250 bpi)<br>- 2 3803-002 Control Units<br>- 1 3803-1792 Two Control Switching Option | |
| • Mass Storage System | 3,374,714 |
| - 472 Billion Bytes (Total)<br>2 3851-A04 Mass Storage Facilities (MSF)<br>236 Billion Bytes (each MSF)<br>4 Data Recording Controls (each MSF)<br>8 Data Recording Devices (each MSF)<br>4720 Cartridges (each MSF) @ $35 each<br>- 2 3830-003 Storage Control Units<br>- 2.536 Billion Bytes Staging Disk<br>(2 3350-A02, 2 3350-B02) | |

TABLE 4.2-3 (Concluded)

COSTS OF A IBM SINGLE MAINFRAME CDHF IMPLEMENTATION

| ITEM | GSA PURCHASE PRICE ($) |
|---|---|
| • Terminals | 62,367 |
| - 1 3274-D31 Control Unit (1 each 6901, 6902 Adapters) | |
| - 10 3278-002 KB/CRT's | |
| - 4 3287-002 Printers (Friction Feed) | |
| • Line Printers | 82,500 |
| - 2 3203-005 (1200 lpm, train cartridge) | |
| • Communications | 86,890 |
| - 1 3705-F04 (24 Bi-sync lines @ 9.6 Kbps) | |
| CDHF TOTAL COST | 9,539,135 |

TABLE 4.2-4

A CDC SINGLE MAINFRAME SUBSYSTEM SUMMARY

| Subsystem | Manufacturer | Comments |
|---|---|---|
| Processor | CDC | Series 700 170-760 |
| Tape | CDC | 4 Drives |
| Disk | CDC | 13.8 Billion 6 bit characters; approximately 50% used for recent production data. |
| Communications | CDC | |
| CDHF Terminals | CDC | 10 CRT Terminals and 4 Printers |
| Array Processor | | None |
| On-Line Mass Storage | CDC | M-860 systems marketed and supported by CDC; expandable |

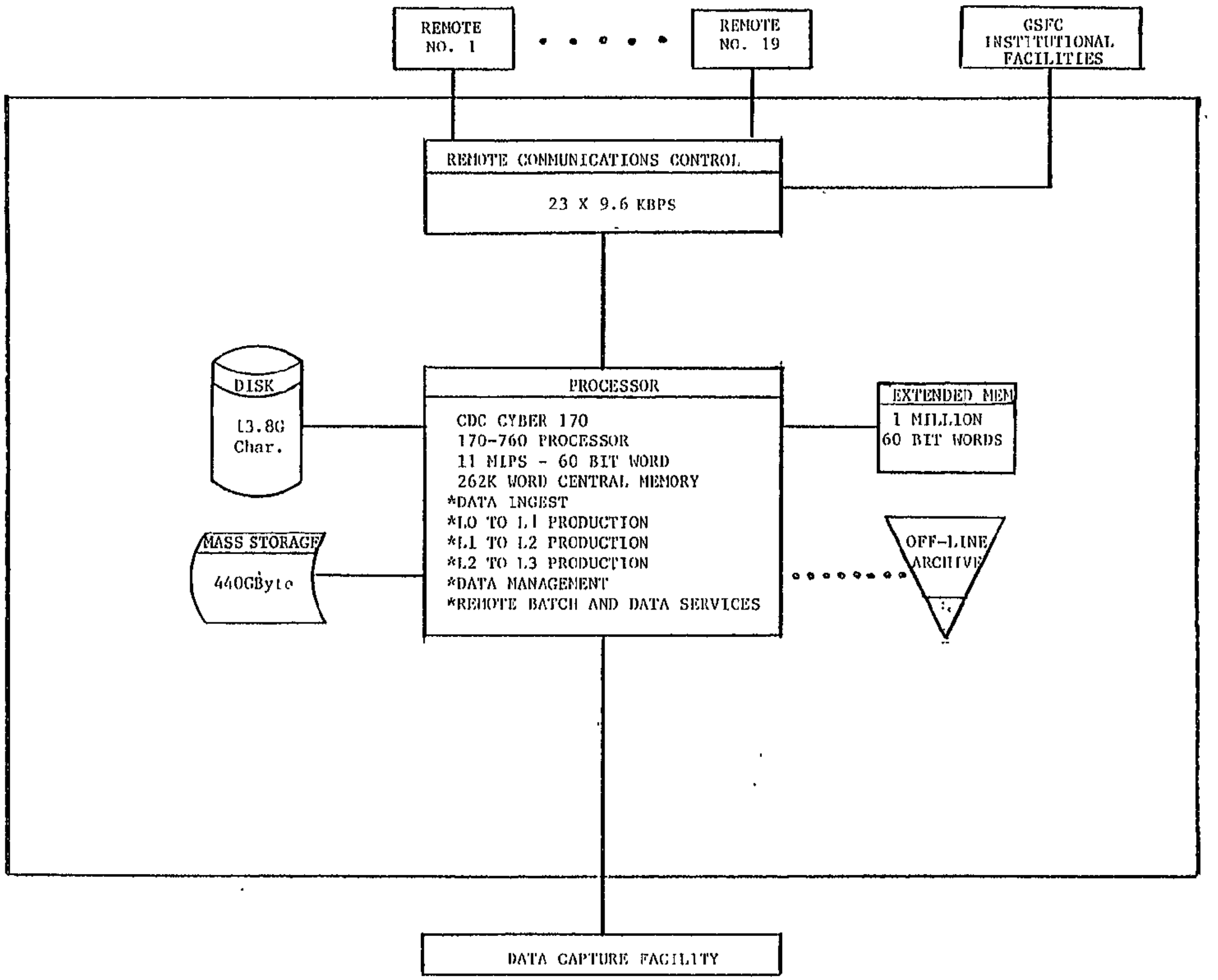
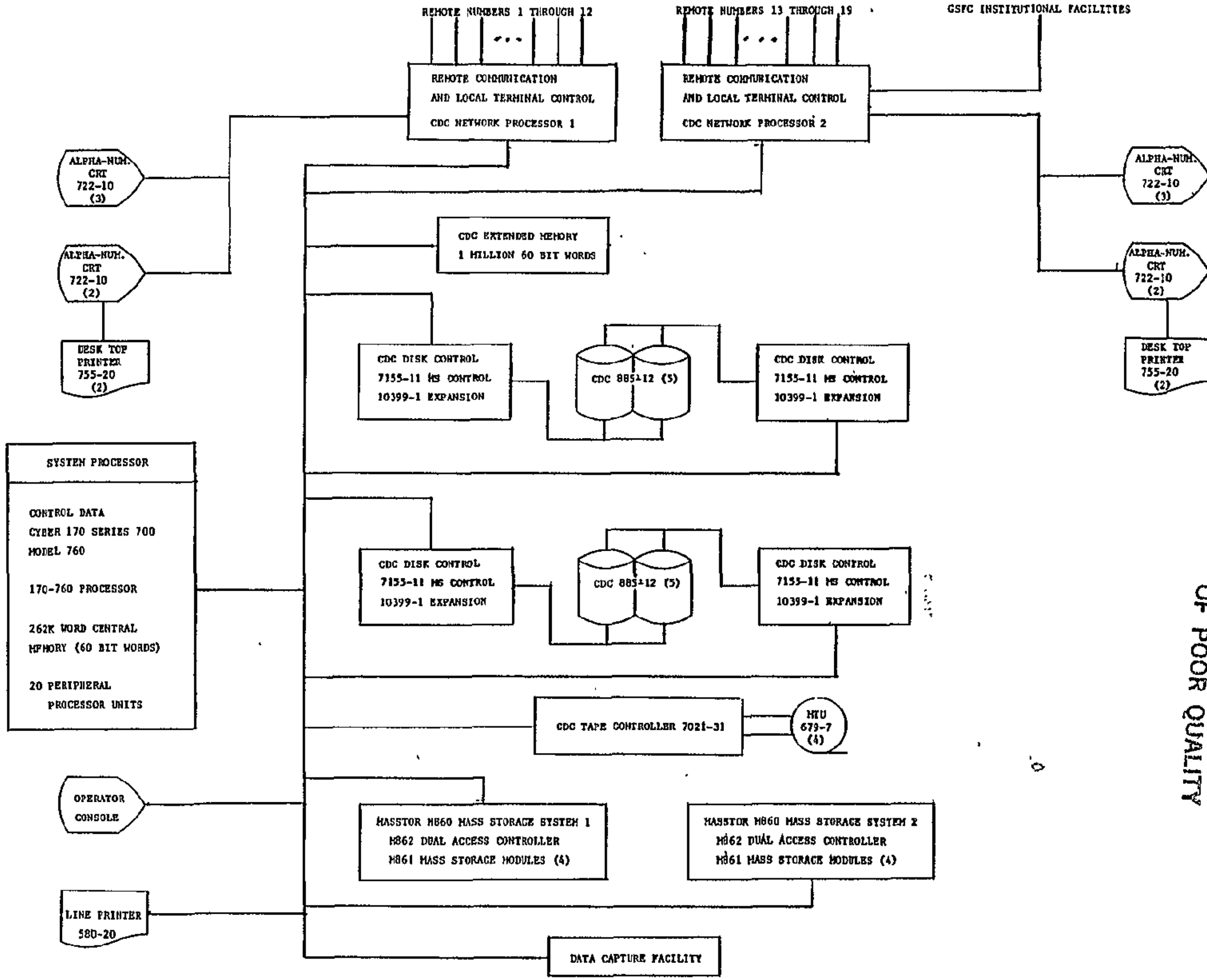



FIGURE 4.2-4 A CDC SINGLE MAINFRAME STRUCTURAL SUMMARY




ORIGINAL PAGE IS OF POOR QUALITY

FIGURE 4.2-5 A CDC SINGLE MAINFRAME IMPLEMENTATION

TABLE 4.2-5

COSTS OF A CDC SINGLE MAINFRAME CDHF IMPLEMENTATION

| ITEM | GSA PURCHASE PRICE ($) |
|---|---|
| • CDC Cyber 170 Series 700<br>170-760 Processor<br>- 11 MIPS<br>- 60 Bit Word<br>- 262K Word Central Memory<br>- 24 Channels<br>- 20 Peripheral Processors<br>- Extended Memory Interface<br>- Power Unit<br>- Operator Console | 3,533,200 |
| • Extended Memory<br>- 1 Million 60 Bit Words<br>- 10 Million Words/Second | 625,000 |
| • Disk (CDC)<br>- 13.84 billion 6-bit characters<br>- 10 885-12 Disk Storage Units (dual spindle, two-controller)<br>- 4 7155-11 Controllers<br>- 4 7155-885 Four Drive Expansion | 855,360 |
| • Tape (CDC)<br>- 4 679-7 Tape Transports (200 ips, 1600/6250 bpi)<br>- 1 7021-31 Controller | 176,340 |
| • Mass Storage System (MASSTOR)<br>- 440 billion bytes<br>- 8 M861 Mass Storage Modules (16 read/write stations)<br>- 2 M862 Dual Access Storage Controllers<br>- 2 Channel Couplers (CDC 65206-X) | 2,793,400 |

TABLE 4.2-5 (Concluded)

COSTS OF A CDC SINGLE MAINFRAME CDHF IMPLEMENTATION

| ITEM | GSA PURCHASE PRICE ($) |
|---|---|
| • 10 Terminals | 26,980 |
| - 10 Alphanumeric CRT's (CDC 722-10)<br>- 4 Desktop Printers (CDC 755-20)<br>- Installation Charges | |
| • Line Printer | 95,078 |
| - 2000 lpm (CDC 580-20)<br>- Printer Train Cartridge (CDC 596-6) | |
| • 2 Communication Subsystems | 131,030 |
| - 2 CDC 2551-1 Network Processing Units (NPU)<br>96K 16-bit words/NPU<br>2 558-3 Couplers<br>6 Sync. Comm Line Adapters (CLA) per NPU (2 Remote Lines/CLA)<br>3 Async. CLA's per NPU (2 Local CRT's/CLA)<br>- Remote Site Interface<br>22 Bi-sync Lines (11/NPU)<br>9600 bps<br>- CRT Terminal Interface (Hardwired)<br>10 Asynchronous Lines<br>9600 bps | |
| CDHF TOTAL COST | 8,236,388 |

## 4.3 Production Processing Demands/Estimates for UARS

A summary of minimum production processing demands on the dual mainframe and single mainframe systems described in Section 4.1 and 4.2 is presented in Table 4.3-1. It should be noted that these processing demands are minimal in that they do not include overhead resources such as those consumed by the operating system.

Table 4.3-2 summarizes the minimal input, output and processing resources that would be consumed by the IBM dual mainframe, CDC dual mainframe, and CDC single mainframe implementation. Again, as was noted for Table 4.3-1, the values listed in Table 4.3-2 are minimal since operating system resource demands and system inefficiencies are not included.

## 4.4 OPEN/UARS CDHF Commonality

Since the on-line storage required for OPEN is about 418 Gbytes (see Table 3.1-4), this is in the range of the mass storage systems envisioned for the UARS CDHF. Thus, major UARS system upgrades are only required to accommodate the higher OPEN processing load. The upgrade could be accomplished as follows: for the dual mainframe approach, the Production Processor (PP) would be substantially upgraded; for the large single mainframe approach, array processors would be added. The latter approach appears to be the more straightforward and appears to offer the greater potential for achieving of hardware and software commonality. An explanation is given in the paragraphs that follow.

Recall that for UARS, it is felt that a computer in the 10-11 megainstructions/sec range could accommodate all UARS processing, with

TABLE 4.3-1(a)

MINIMUM PROCESSING DEMANDS ON IBM DUAL MAINFRAME

| CATEGORY | | DM/P (0.75 MIP) | | PP (3.0 MIP) | | | |
|---|---|---|---|---|---|---|---|
| | | IBM 4341-L02 | % of 8 hours | IBM 3033-N-8 | % of 8 hours | AP190L | AP190L |
| INGEST [1] | | 2133 sec | 7.4 | | | | |
| L-1 | | 8116 sec | 28.2 | | | | |
| L-2 | | | | 8580 | 29.8 | 11400 | 11900 |
| L-3 | | 3963 sec | 13.8 | | | | |
| TOTALS: | SECONDS | 14212 | 49.3 | 8580 | 29.8 | 11400 | 11900 |
| | HOURS | 3.948 | | 2.38 | | 3.16 | 3.305 |

<u>Note</u>: (Applicable to all of this table)

[1] Mapping from raw data to L-0 assumes 8 instructions/L-0 byte; does not include checksum processing.

ORIGINAL PAGE IS
OF POOR QUALITY

TABLE 4.3-1(b)

MINIMUM PROCESSING DEMANDS ON CDC DUAL MAINFRAME

| CATEGORY | | DM/P (1.2 MIP) | | PP (3.5 MIP) | | |
|---|---|---|---|---|---|---|
| | | CDC 170-720 | % of 8 hours | CDC 170-730 | % of 8 hours | Advanced Flexible Processor |
| INGEST [1] | | 1333 | 4.6 | | | |
| L-1 | | 5073 | 17.6 | | | |
| L-2 | | | | 7354 | 25.5 | 2807 [2] |
| L-3 | | 2477 | 8.6 | | | |
| TOTALS: | SECONDS | 8883 | 30.8 | 7354 | 25.5 | 2807 |
| | HOURS | 2.468 | | 2.043 | | 0.779 |

<u>Note</u>:

[1] Mapping from raw data to L-0 assumes 8 instructions/L-0 byte; does not include checksum processing.

[2] Advanced Flexible Processor performs up to 200M arithmetic operations/second; 1/8 rate assumed in these estimates; must be coded in assembly language.

TABLE 4.3-1(c)

MINIMUM PROCESSING DEMANDS ON IBM SINGLE MAINFRAME

| CATEGORY | IBM 3081 (10.4MIP) [3] | % of 8 hours |
|---|---|---|
| INGEST | 154 | 0.5 |
| L-1 | 585 | 2.0 |
| L-2 | 9223 | 32.0 |
| L-3 | 286 | 1.0 |
| TOTALS: SECONDS | 10248 | 35.6 |
| TOTALS: HOURS | 2.847 | |

Note:

[3] No external array processor.

TABLE 4.3-1(d)

MINIMUM PROCESSING DEMANDS ON CDC SINGLE MAINFRAME

| CATEGORY | CDC 170-760 (11 MIP) [3] | % of 8 hours |
|---|---|---|
| INGEST | 145 | 0.5 |
| L-1 | 553 | 1.9 |
| L-2 | 8720 | 30.3 |
| L-3 | 270 | 0.9 |
| TOTALS: SECONDS | 9688 | 33.6 |
| TOTALS: HOURS | 2.691 | |

Note:

[3] No external array processor.

TABLE 4.3-2

MINIMUM PRODUCTIVE RESOURCE DEMANDS (HOURS)

| CATEGORY | | IBM DUAL MAINFRAME | | CDC DUAL MAINFRAME | | IBM SINGLE MAINFRAME | CDC SINGLE MAINFRAME |
|---|---|---|---|---|---|---|---|
| | | DM/P | PP | DM/P | PP | | |
| INGEST: | READ RAW | 0.79 | - | 0.79 | - | 0.79 | 0.79 |
| | PROCESS | 0.59 | - | 0.37 | - | 0.04 | 0.04 |
| | WRITE L-0 | 0.56 | - | 0.56 | - | 0.56 | 0.56 |
| L-1 PROC: | READ L-0 | 0.56 | - | 0.56 | - | 0.56 | 0.56 |
| | PROCESS | 2.25 | - | 1.41 | - | 0.16 | 0.15 |
| | WRITE L-1 | 1.00 | - | 1.00 | - | 1.00 | 1.00 |
| L-2 PROC: | READ L-1 | - | 1.00 | - | 1.00 | 1.00 | 1.00 |
| | PROCESS | - | 2.38[1] | - | 2.04[2] | 2.56 | 2.42 |
| | WRITE L-2 | - | 0.37 | - | 0.37 | 0.37 | 0.37 |
| L-3 PROC: | READ L-2 | 0.37 | - | 0.37 | - | 0.37 | 0.37 |
| | PROCESS | 1.10 | - | 0.69 | - | 0.08 | 0.08 |
| | WRITE L-3 | 0.13 | - | 0.13 | - | 0.13 | 0.13 |
| | INPUT[3] | 1.72 | 1.00 | 1.72 | 1.00 | 2.72 | 2.72 |
| TOTALS: | PROCESS | 3.94 | 2.38 | 2.47 | 2.04 | 2.84 | 2.69 |
| | OUTPUT[3] | 1.69 | 0.37 | 1.69 | 0.37 | 2.06 | 2.06 |

Note:

[1] In addition, concurrent array processing consumes 3.16 hours of AP190L(1) resources and 3.31 hours of AP190L(2) resources.

[2] In addition, concurrent array processing consumes 0.78 hours of Advanced Flexible Processor (AFP) resources.

[3] "Ingest; READ RAW" estimate based on 1.4619 x $10^9$ bits input at 512K bits/sec; all other "READ/WRITE" estimates use an estimate of 10 usec per byte.

no attached array processor required, in about 3 hours (theoretical throughput). Since the OPEN processing load is estimated to be about 5.2 times that of UARS (Section 3.2.2), it would appear that about 15.6 hours of the UARS mainframe would be required for OPEN. However, five of the OPEN instruments are estimated to consume 90% of the OPEN processing load (also Section 3.2.2), and these instruments' data are suitable for efficient array processing. Hence, the attachment of array processors to the UARS single mainframe offers promise for significantly reducing the 15.2 hour demand on the computer. If this is the case, without substantial hardware design, commonality could be achieved.

In addition to the common hardware design inherent in this approach, there could be promise for achieving a measure of software commonality. As seen in Section 5, there appear to be substantial areas of commonality between the OPEN and UARS software systems both in the areas of data management software and the production software. If both OPEN and UARS processing utilized the same mainframe, then the software would be available to both and substantial cost savings could be realized.

# 5.0 CDHF DATA PROCESSING AND MANAGEMENT CONCEPT

## 5.1 Introduction

Given that a UARS or an OPEN CDHF can be configured from readily available commercial hardware subsystems such as those presented in Section 4, this section will point out several challenges which require in depth exploration in order to provide remote users with rapid and reliable access to the massive volumes of UARS and OPEN data which will be stored on-line at the respective CDHF's. The quantity of on-line data at the UARS CDHF will be in the neighborhood of 200 gigabytes. OPEN CDHF on-line data requirements, being in excess of 400 gigabytes, are more than double those of UARS.

UARS investigators have submitted preliminary lists of retrieval keys and browsing criteria, as listed in Tables 5.0-1, 5.0-2, and 5.0-3. OPEN investigators are expected to submit similar requirements in the future. An implication of the lists of retrieval keys and browsing criteria listed in Tables 5.0-1, 5.0-2, and 5.0-3 is that use of a data base management system (DBMS) and query language might be employed at the respective CDHF's to facilitate the remote user interface with the on-line data. However, the decision to employ a DBMS, such as today's commercially available products, must be carefully investigated. In particular, the topics of access speed, data base recovery and data base reorganization must be considered.

Adequate data access speeds in a DBMS environment involving data bases of several hundreds of gigabytes could require that extensive additional high speed disk facilities be added to the hardware con-

| | TIME OR UT | GEOGRAPHIC LOCATION/LAT/LONG | MODE OF OPERATION | INSTRUMENT | DAY-NIGHT | POINTED PLATFORM COORDS (RIGHT ASC & DEC) | GEOMAGNETIC LOCATION | ALTITUDE | AE INDEX | Kp INDEX | $D_{st}$ INDEX | SOLAR FLUX | S/C ALTITUDE | "SPECIES FOR HIGHER LEVELS" | ORBIT NUMBER | TAPE RECORDER PLAYBACK NUMBER | SENSOR DETAILS | DATE AND TIME OF CREATION | GENERATION NUMBER | ALGORITHM VERSION NUMBER | UARS YAW POSITION | MISSION ELAPSED TIME | S/C ATTITUDE | S/C PLATFORM POINTING VECTOR | S/C LINE OF SIGHT IMACT ALTITUDE | INSTRUMENT OPERATIONAL MODE/DATA SET TYPE | IRRADIANCE AND VARIABILITY INDICES | TBD |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| WINTERS | ● | ● | ● | | | | | | | | | | | | | | | | | | | | | | | | | |
| HRDI | UNKNOWN | | | | | | | | | | | | | | | | | | | | | | | | | | | ● |
| CLAES | UNKNOWN | | | | | | | | | | | | | | | | | | | | | | | | | | | ● |
| HALOE | ● | ● | | | | | | | | | | | | ● | | | | | | | | | | | | | | |
| ISAMS | ● | ● | | | | | | | | | | | | | ● | ● | ● | ● | ● | ● | | | | | | | | |
| MLS | ● | ● | | | | | | | | | | | | | | | | | | | ● | | | | | | | |
| PEM | ● | ● | | | | | ● | ● | ● | ● | ● | ● | ● | | | | | | | | | | | | | | | |
| SUSIM | | ● | | | | | | | | | | | | | ● | | | | | | | ● | ● | ● | ● | ● | ● | |
| SOLSTICE | ● | | | ● | ● | ● | | | | | | | | | | | | | | | | | | | | | | |
| SBUV | UNKNOWN | | | | | | | | | | | | | | | | | | | | | | | | | | | ● |

TABLE 5.0-1 UARS PRIMARY RETRIEVAL KEYS

| | EXTREME VALUES (INTENSITY, TEMP., ETC.) | GEOGRAPHIC LOCATION/LAT/LONG | WAVELENGTH | DATA QUALITY | COINCIDENCE WITH CORRELATIVE MEASUREMENTS | PROCESSING STATUS | INSTRUMENT STATUS (OWN/OTHER) | ATMOSPHERIC EVENTS | AURORAL EVENTS, ETC. | SOLAR EVENTS | SPACECRAFT STATUS | ALGORITHM CONFIDENCE CODES | SPECIAL EVENTS | SOLAR ACTIVITY | TBD | TO BE SPECIFIED BY OPS PERSONNEL (ACCORDING TO EXPERIMENT) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| WINTERS | • | | | | | | | | | | | | | | | |
| HRDI | UNKNOWN | | | | | | | | | | | | | | | |
| CLAES | UNKNOWN | | | | | | | | | | | | | | | |
| HALOE | | | | • | • | • | | | | | | | | | | |
| ISAMS | | | | • | | | • | • | • | • | • | • | | | | |
| MLS | | | | | • | | | | | | | | • | | | |
| PEM | | | | | | | | | | | | | | | • | |
| SUSIM | | | | | | | | | | | | | | • | | • |
| SOLSTICE | | • | • | | | | | | | | | | | | | |
| SBUV | UNKNOWN | | | | | | | | | | | | | | | |

TABLE 5.0-2 UARS SECONDARY RETRIEVAL KEYS

ORIGINAL PAGE IS OF POOR QUALITY

| | TEMPERATURE AT SPECIFIC ALTITUDE | WIND (2 COMPONENTS) AT SPECIFIC ALTITUDE | INSTRUMENT | DATE AND TIME | TARGET | WAVELENGTH | DATA QUALITY/DATA VALIDATION LEVEL | ORBIT NUMBER | BEGIN/END GLOBAL POSITION FOR ORBITAL SEG. | IONIZING RADIATION INDEX AND LOCATION INDICATOR AT DISCRETE ALTITUDE INTERVALS | EVIDENCE OF MINOR SPECIES CONCENTRATION CHANGE | EVIDENCE OF UNUSUAL ENHANCEMENT OF SCATTERED SUNLIGHT | GEOGRAPHIC LOCATION | DATATYPE - INSTRUMENT | - MODULATIONS | - SPECIES PROFILES/BURDENS | DATA COVERAGE - EVENT | - GRIDDED | - SPACE/TIME AVERAGED |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| WINTERS | • | • | | | | | | | | | | | | | | | | | |
| HRDI | UNKNOWN | | | | | | | | | | | | | | | | | | |
| CLAES | | | | • | | | | • | • | | | | | | | | | | |
| HALOE | | | | • | | | • | | | | | | • | • | • | • | • | • | • |
| ISAMS | | | | • | | | • | • | | | | | • | | | | | | |
| MLS | | | | | | | • | | | | | | | | | | | | |
| PEM | | | | | | | | | | • | • | • | | | | | | | |
| SUSIM | | | | | | | | | | | | | | | | | | | |
| SOLSTICE | | | • | • | • | • | • | | | | | | | | | | | | |
| SBUV | UNKNOWN | | | | | | | | | | | | | | | | | | |

| | CORRELATIVE MEASUREMENT MATCH CODE | PROCESSING LEVEL | PEM MAGNETIC DISTURBANCE SUMMARIES (TIME, POS, INTENSITY) | ICE MAGNETIC DISTURBANCE SUMMARIES (NOT AVAIL.) | VARIOUS GEOPHYSICAL PARAMETERS | ALTITUDE RANGE COVERED | TAPE RECORDER PLAYBACK NUMBER | SENSOR DETAILS | DATE AND TIME OF CREATION | GENERATION NUMBER | ALGORITHM VERSION NUMBER | INSTRUMENT STATUS (OWN/OTHER) | ATMOSPHERIC EVENTS | AURORAL EVENTS, ETC. | SOLAR EVENTS | SPACECRAFT STATUS | ALGORITHM CONFIDENCE CODES | OTHERS TBD |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| WINTERS | | | | | | | | | | | | | | | | | | |
| HRDI | | | | | | | | | | | | | | | | | | |
| CLAES | | | | | | | | | | | | | | | | | | |
| HALOE | • | • | | | | | | | | | | | | | | | | |
| ISAMS | | | | | | | • | • | • | • | • | • | • | • | • | • | • | |
| MLS | | | | | • | • | | | | | | | | | | | | |
| PEM | | | | | | | | | | | | | | | | | | • |
| SUSIM | | | | | | | | | | | | | | | | | | • |
| SOLSTICE | | | | | | | | | | | | | | | | | | |
| SBUV | | | | | | | | | | | | | | | | | | |

TABLE 5.0-3 UARS CATALOG/BROWSE ENTRIES

figurations presented in Section 4. This could significantly increase the costs of the configurations presented in Section 4 since all of the configurations rely on relatively inexpensive IBM or MASSTOR mass storage facilities (MSF's) to minimize the costs of on-line data storage. MSF costs are approximately $70,000 for each 10 gigabytes of storage, while a corresponding amount of high speed disk storage would cost in excess of $1.5 million. Both the IBM and MASSTOR devices store data on randomly accessible magnetic tape strips. Access times for these devices are approximately five to ten seconds per tape strip and can be even longer when all available read/write stations are in use. The ability to configure these devices into a DBMS system is essential if CDHF storage costs are to be kept at a minimum.

Any DBMS considered for use at the CDHF must possess features which minimize the amount of the data that must be restored following data-destructive hardware or software malfunctions. Halfway through the lifetime of the UARS CDHF complete restoration of the on-line data from backup tapes would involve reading several thousand reels of tape over a period of several hundreds of hours. The data base restoration capabilities of any DBMS considered for the UARS or OPEN CDHF must be among the prime considerations.

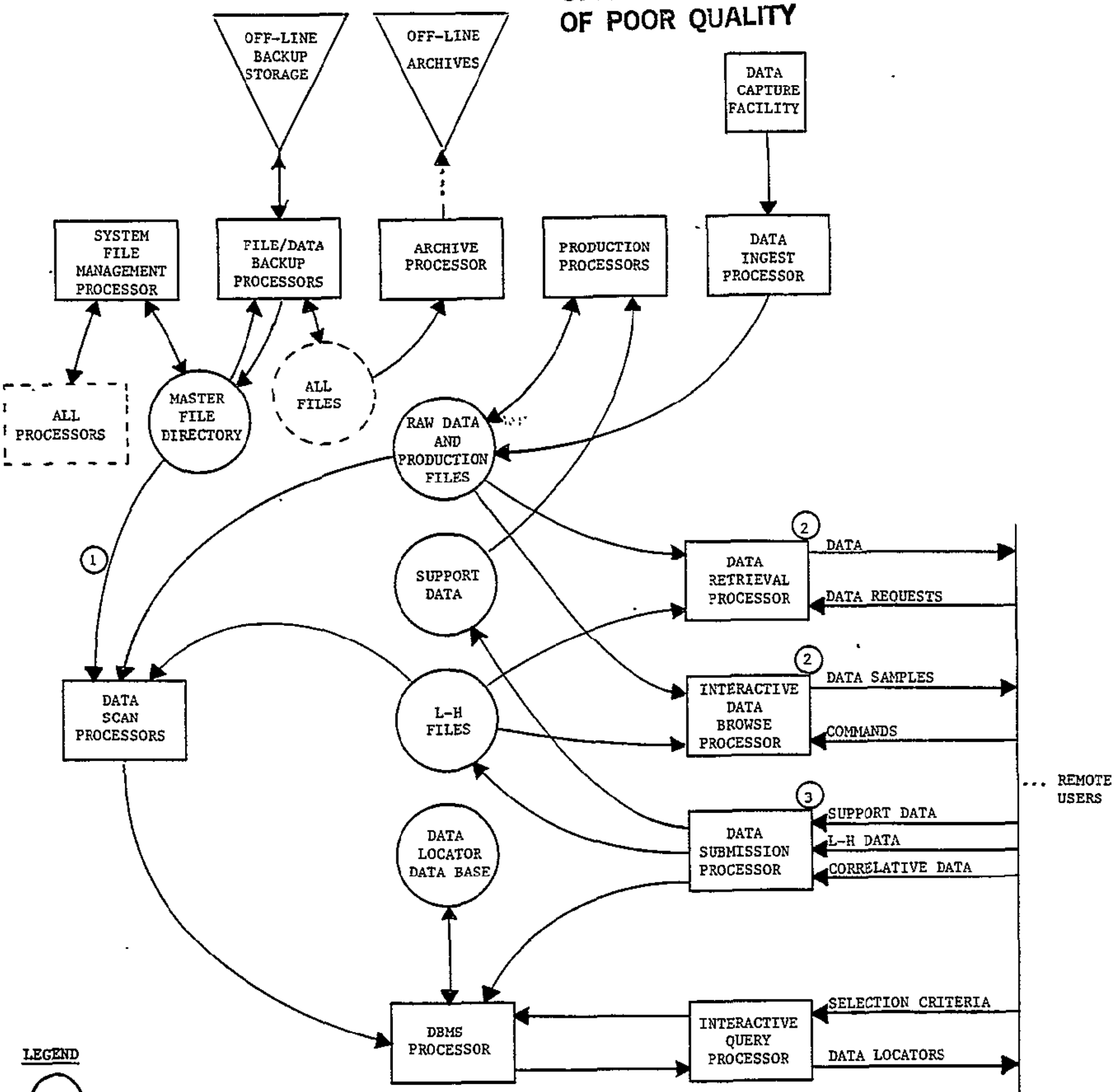
Any DBMS system considered for use at the UARS or OPEN CDHF must possess features which will minimize the time and effort required to reorganize the data base in order to overcome unacceptable access speed deterioration or potential storage saturation. Once significant amounts of data have been collected reorganization of the data base

would become impractical if it involved the re-entry of extensive amounts of data from several thousand reels of tape.

The remaining paragraphs of this section present a unified data processing and management concept which would simplify the data restoration process and eliminate the possibility of having to reorganize massive data bases.

## 5.2 Data Processing and Management Concept

A description of a common approach to UARS and OPEN data management is provided in the following paragraphs and summarized pictorially in Figure 5.2-1. The concept presented would make use of standard system sequential file processors (vendor utilities) to manage the large quantities of UARS or OPEN data at the CDHF and would thus minimize storage overhead for these data. Use of sequential files would also simplify any data restoration process required to recover data destroyed as the result of system malfunctions. Since large sequential files do not lend themselves to rapid querying by remote users, a Data Locator Data Base (DLDB) designed for fast access would be provided to assist users in locating data of interest. The DLDB would be event and condition oriented and would be of a coarser time granularity than the UARS and OPEN data that is summarized. Such a data base, being a summary of the data elements used to derive it, could be much more compact and rapidly accessible than would be a data base which consisted of the constituent data elements of the events themselves. Vectors into specific files that contained data




LEGEND

○ = ON-LINE DATA FILES OR DATA BASE

□ = SOFTWARE PROCESSORS

▽ = OFF-LINE STORAGE

NOTES:
1. FILE ATTRIBUTES (TIME OF CREATION, LAST MODIFICATION, ETC.)
2. "READ ONLY" ACCESS.
3. ANY USER MAY SUBMIT DATA, A USER MAY ONLY UPDATE/DELETE HIS OWN DATA.

FIGURE 5.6-1

CDHF DATA PROCESSING AND MANAGEMENT CONCEPT

corresponding to particular events or conditions would be filed in the DLDB to provide the user with the information necessary to access, browse through or retrieve data of interest.

### 5.2.1 Production Cycle

Similarities between UARS and OPEN suggest that a common software production framework and collection of production software utilities could be configured for use at both CDHF's.

As pictured in Figure 5.2-1, a typical production cycle would begin with the arrival of spacecraft data via the data capture facility. Raw data would be read into the CDHF data processing system by the Data Ingest Processor, subjected to elementary quality control checks, and stored. Subsequently the Production Processors (PP) would be activated in turn. These production processors need not be unique to a particular CDHF. During the production processing the various PP's would input raw or L-(n) data and any required spacecraft or instrument oriented support data and produce a specific L-(n+1) output. As various segments of the production process are completed, appropriate Data Scan Processors (DSCAN) would be activated. The function of the DSCAN processors would be to examine the various new production files registered in the system Master File Directory (MFD) and to develop (predefined) summary information for incorporation into the Data Locator Data Base (DLDB). As an example, the DSCAN might note at which point in time a peak reading occurred in a particular subsystem and record such items as (1) the value of the reading, (2)

the (real) time and date at which the event occurred, (3) instrument status, (4) the name of the data file containing the reading, and (5) the relative position (within the data file) of the reading. Subsequently the DSCAN would submit this (and other) significant events data to the Data Base Management System Processor (DBMSP) for incorporation into the DLDB. Once incorporated into the DLDB, the significant events and vectors into the production data file system would be available to the user community via the Interactive Query Processor (IQP).

### 5.2.2 User Interface

The concept presented in Figure 5.2-1 provides four processors to allow users to submit, locate and retrieve data. These processors are as follows:

- Data Submission Processor (DSUB)
- Interactive Query Processor (IQP)
- Interactive Browse Processor (IBP)
- Data Retrieval Processor (DRP)

Table 5.2-1 illustrates how these and other CDHF processors could be used to conduct activities analogous to those carried out at a conventional library of printed books.

#### 5.2.2.1 Data Submission Processor

The DSUB processor would permit users to submit a variety of data into the CDHF data system. Three major types of data (correlative, support and L-H data) are noted in Figure 5.2-1. Incoming correlative

TABLE 5.2-1. CDHF VERSUS CONVENTIONAL LIBRARY FUNCTIONS

| Function | Conventional Library | CDHF |
|---|---|---|
| Data Generation | Book purchases. | Production Programs generate L-0, L-1, L-2 and L-3 data. Remote users submit higher level L-H data (>L-3) and correlative data using the Data Submission Processor. |
| Indexing | Index card files are updated with indications of new volumes and subject matter. | Data Scan Processor examine new or changed L-0, L-1, L-2, L-3 and L-H data and generate updates for the Data Locator Data Base. Correlative data are used to update the Data Locator Data Base. |
| Subject Matter Search | User examines index card files and notes volume numbers and pages of interest. | User accesses Data Locator Data Base using the Interactive Query Processor and is presented with lists of files and locations within files wherein might reside subject matter of interest. |
| Text Search | User requests volumes of interest and browses pages of interest; specific data of interest are determined. | User accesses potential areas of interest within data files using the Interactive Browse Processor; user determines data of interest. |
| Volume/Text/Data Acquisition | User checks out volume of interest or copies pages of interest. | User instructs Data Retrieval Processor to transmit copies of data sets or subsets to user's remote facility. |

data, such as measurements from ground observations or sounding rockets, would be forwarded to the DBMS processor for incorporation into the DLDB. Support data, such as instrument calibration data, would be vectored into Support Data Files (SDF) where they would become available to the various PP's. Incoming L-H data (levels of processing greater than standard production data) developed at remote user facilities would be vectored into L-H Files (L-HF). Newly received or updated L-HF data would be scanned by appropriate DSCAN processors for significant events and the DLDB would be updated (using the DBMS processor) in a manner similar to that used with the production files.

#### 5.2.2.2 Interactive Query Processor

The IQP, operating in conjunction with the DLDB, would be a key user/system interface. PI's have already indicated how they will desire to locate data residing at the CDHF; Tables 5.0-1, 5.0-2 and 5.0-3 summarize the data attributes which UARS investigators identified as being of interest. In the concept presented in Figure 5.2-1 the data attributes listed in Tables 5.0-1, 5.0-2 and 5.0-3 (along with others to be defined) would be used to formulate queries which would be submitted to the IQP. Various attributes, such as those listed in the tables, would be linked together logically with delimiting values to form queries describing the data of potential interest. Figure 5.2-2 illustrates how such a query might appear. The data presented as the response in Figure 5.2-2 would have been previously

QUERY: LOCATE DATA FOR SUBSYSTEM name:

CRITERIA ARE:

ORBIT-NUMBER = m or n

AND DATA-QUALITY is xxxx

AND IRRADIANCE-INDEX $\geq$ yyy;

LIST DATA-SET-NAME, DATA-LOCATOR-NUMBER.

RESPONSE:

| DATA-SET-NAME | DATA-LOCATOR-NUMBER |
|---|---|
| name1 | 115-120 |
| | 135-166 |
| | 168 |
| name2 | 350 |
| | 355-360 |
| | 363 |

FIGURE 5.2-2. SAMPLE QUERY AND RESPONSE

entered into the DLDB by the DSCAN processors previously discussed. The data set names and data set numbers could subsequently (by either automatic or manual interfaces) be used as inputs to drive the Interactive Browse Processor and the Data Retrieval Processor.

#### 5.2.2.3 Interactive Browse Processor

The purpose of the IBP would be to permit remote users to visually examine selected data fields within a particular production or L-H file. Use of the IBP would usually be preceded by use of the IQP to determine the general location(s) of data of interest. Once the general location(s) of the data of interest had been determined the user would instruct the IBP to position the data file of interest to the general area of interest. Subsequently, the user would command the IBP to position the file forward and/or backward and to display specific data fields of interest as illustrated in Figure 5.2-3.

#### 5.2.2.4 Data Retrieval Processor

Remote users would use this processor to designate sets or subsets of production or L-H data for which copies are desired. The Data Retrieval Processor (DRP), in turn, would locate the specific data and transmit a copy of that data to the remote user. Use of the DRP would frequently be the final activity in a QUERY-BROWSE-RETRIEVE sequence of activity.

### 5.2.3 Data Security

Since the data stored at the CDHF will represent significant expenditures of manpower, analytic and data processing resources, it

```
COMMANDS:  OPEN filename

           ADVANCE TO DATA-LOCATOR-NUMBER

           DISPLAY name1, name2, name3, GMT

RESPONSE:  name1 = data1

           name2 = data2

           name3 = data3

             GMT = time

COMMANDS:  ADVANCE UNTIL GMT = time1

           DISPLAY name2, GMT

RESPONSE:  name2 = data2

             GMT = time1
                 .
                 .
                 .
```

FIGURE 5.2-3. BROWSE SEQUENCE

is essential that the CDHF include the necessary elements of security to protect data from accidential or deliberate loss or destruction. The concept presented in Figure 5.2-1 would include provisions for data security that would minimize the chances for the loss or destruction of data by providing off-line backup copies of data and by controlling accesses to on-line data.

5.2.3.1 Off-Line-Backup

In the concept presented in Figure 5.2-1 the creation of off-line backup copies of on-line data would be provided using File/Data Backup Processors (F/DBP). While processors could be especially written for the CDHF, in many cases off-the-shelf system utilities are available to provide backup file copies on magnetic tape.

The creation of off-line backup data copies (probably using magnetic tape as a backup medium) would be an ongoing process throughout the lifetime of the CDHF. F/DBP's, working in conjunction with an overall System File Management Processor (SFMP) could insure that backup copies of all newly established or modified on-line data would be created on a regular basis. In the event of system software or hardware failures or user errors resulting in the loss of on-line data, an appropriate F/DBP could re-establish the data on-line from the most recent backup copy available for that data. Such a recovery method would avoid the necessity of lengthly (multi-hour, in some cases) computer runs to recreate lost data from lower level data. Figure 5.2-1 illustrates a flow of data between on-line storage and

local off-line storage. While the local off-line storage facility would be a functional element of the CDHF, the physical location of the storage facility itself might be somewhat removed from the CDHF computing facility in order to preclude the loss of both on-line and off-line copies of data in the event of a catastrophic event such as a fire in the CDHF.

#### 5.2.3.2 Data Access Controls

Since the CDHF system will be a multi-user system, access to on-line data will of necessity be limited to a certain degree (at a minimum it would be necessary to prohibit concurrent updating of the same data by multiple users and/or processors). An indication of the types of user access control that will be required is as follows:

- Any user or appropriate processor may gain read access to any data which is not being written, modified or deleted by another user or processor.
- Production data may be created, modified or deleted only by the Production Processors, the File/Data Back Processors or the CDHF staff.
- Any user may submit new correlative data, support data or L-H data, subject to predefined authorization restrictions.
- Any user may augment, modify or delete only the data that he has created and, further, only when such activities would not conflict with the access of another user or processor.

In terms of Figure 5.2-1, these types of access control could be provided by the SFMP and the DBMS processor. User identification codes or account numbers augmented (if necessary) by "add/modify/delete" privilege passwords would probably prove adequate.

5.2.4 Archive Function

Details regarding the extensive quantities of data that must be archived will be presented in Section 6. Depending upon the nature and compatibility of the archive medium and/or interface, the Archive Processor included in Figure 5.2-1 would prepare (1) archive data using the actual archive medium and format or (2) intermediate copies of archive data, using a medium such as a computer compatible tape, for subsequent transcription onto the archive medium. While archive medium generation for production and higher level data could be postponed until the end of the CDHF lifetime, consideration should be given to an ongoing archive process (perhaps a daily or weekly archive generation run) throughout the CDHF lifetime. An ongoing archive process (of data which have become static) could preclude the necessity for an extended series of archive production runs involving the transfer of hundreds of billions of bytes of data from on-line storage. The introduction of the concept of an ongoing archive process might also lead to significant savings in time and resources if it proved feasible to combine certain elements of the archive process and the ongoing data backup process described in Section 5.2.3.1.

# 6.0 DATA ARCHIVES

## 6.0 DATA ARCHIVES

The analysis presented in this section describes archive requirements for the CDHF data. Although the details are derived for UARS data, the UARS numbers can be scaled up to the OPEN numbers keeping in mind the following: Daily OPEN data volume exceeds UARS data volume by about 50%; and total OPEN mission data exceeds total UARS mission data by a factor of about 7 (see Tables 6.1-1 and 3.1-4, respectively).

### 6.1 Data Volume

For UARS, the five types of data discussed in this analysis fall into two general categories:

- Standard data products (L-0, L-1, L-2, L-3)
- Non-standard data products (L-H); i.e., higher order data (>L-3) submitted from remote sites.

Table 6.1-1 lists the estimated volume for each level of UARS data. The standard data products are those produced on a daily basis at the CDHF. Standard data products will be produced at a known rate and the size of each product will be known. The production rates and sizes of the non-standard data products, in contrast, will vary as a function of UARS PI findings and activity. Note that other data at the CDHF such as software, support data (calibration processing coefficients, ground truth measurements, other correlative measurements), ephemeris data, etc., are not considered in this analysis of archive requirements.

It should be noted that the "Megabyte" estimates given for various types of data refer to the space that would be occupied by binary

TABLE 6.1-1. UARS ESTIMATED DATA VOLUMES (Megabytes)

| Data Type | Daily Volume | Total (540 Days) |
|---|---|---|
| L-0 | 199.95 | 107,973 |
| L-1 | 356.70 | 192,618 |
| L-2 | 133.33 | 71,998 |
| L-3 | 48.35 | 26,109 |
| L-H | 16.79 | 9,067 |
| Total | 755.12 MB | 407,765 MB |

Note: Data volumes in terms of space required to store binary images of integer, single precision floating point and double precision floating point values.

images of integer values, single precision floating point values and double precision floating point values; the estimates do not refer to data stored in character format. Conversion of binary data to numeric character strings, as illustrated in Figure 6.1-1, could easily double, triple or quadruple the size of the UARS data archives.

Table 6.1-2 indicates daily and 540 day (total) UARS archive data in terms of binary data stored on 9 track 6250 bpi computer compatible tape (CCT) reels. The information in Table 6.1-2 is based upon the tape capacity data listed in Table 6.1-3. The first part of Table 6.1-2 assumes a daily archive production run in which all archive data are copied to tape sequentially, leaving only the last tape of a daily composite archive set partially filled. If, however, a separate set of archive tapes is produced for each of the five levels of data on a daily basis, the daily volume of archive tapes would increase from the range of 5-7 reels to the range of 8-10 reels as shown in the second part of Table 6.1-2. Likewise, total archive storage would increase from the range of 4500-6300 reels to the range of 7200-9000 reels.

## 6.2 Further Considerations

If CCT in the binary format is chosen as the UARS data archive medium, the quantity of 6250 bpi tape reels that will be generated will be in the range of 4500 to 9000 reels of tape.

Binary image format offers a distinct advantage because of its compactness. However, a distinct disadvantage of storing binary data is that such data will have to undergo extensive pre-processing when

| Byte 1 | Byte 2 | Byte 3 | Byte 4 |
|---|---|---|---|
| 00000000 | 00001111 | 01000010 | 01000000 |

(a) $+1000000_{10}$ stored as 32 bit binary signed integer

| Byte 1 | Byte 2 | Byte 3 | Byte 4 | Byte 5 | Byte 6 | Byte 7 | Byte 8 |
|---|---|---|---|---|---|---|---|
| 00101011 | 00110001 | 00110000 | 00110000 | 00110000 | 00110000 | 00110000 | 00110000 |
| + | 1 | 0 | 0 | 0 | 0 | 0 | 0 |

(b) $+1000000_{10}$ stored as an ASCII character string

FIGURE 6.1-1. AN ILLUSTRATION FOR STORING BINARY VERSUS CHARACTERS

TABLE 6.1-2. UARS 6250 BPI ARCHIVE TAPE VOLUME

| | | COMPOSITE SETS | | | | | SEPARATE SETS | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Data Type | Daily Volume (Megabytes) | Number of 6250 bpi Archive Tapes When Using Record Lengths of (in bytes): | | | | | Number of 6250 bpi Archive Tapes When Using Record Lengths of (in bytes): | | | | |
| | | 4096 | 8192 | 16384 | 32768 | 65536 | 4096 | 8192 | 16384 | 32768 | 65536 |
| L-0 | 199.95 | 1.70 | 1.42 | 1.29 | 1.23 | 1.19 | 2 | 2 | 2 | 2 | 2 |
| L-1 | 356.70 | 3.03 | 2.54 | 2.30 | 2.19 | 2.13 | 4 | 3 | 3 | 3 | 3 |
| L-2 | 133.33 | 1.13 | 0.95 | 0.86 | 0.82 | 0.80 | 2 | 1 | 1 | 1 | 1 |
| L-3 | 48.35 | 0.41 | 0.34 | 0.31 | 0.30 | 0.29 | 1 | 1 | 1 | 1 | 1 |
| L-H | 16.79 | 0.14 | 0.12 | 0.11 | 0.10 | 0.10 | 1 | 1 | 1 | 1 | 1 |
| Total | 755.12 | 6.41 | 5.38 | 4.88 | 4.63 | 4.51 | | | | | |
| Whole Reels Per Day | | 7 | 6 | 5 | 5 | 5 | 10 | 8 | 8 | 8 | 8 |
| Whole Reels For 540 Days | | 3780 | 3240 | 2700 | 2700 | 2700 | 5400 | 4320 | 4320 | 4320 | 4320 |

TABLE 6.1-3. CCT CAPACITY AT 6250 BITS/INCH

| Data Record Size (Bytes) | Data Record Length (Inches)[1] | Records Per 2400 Ft Reel[2] | Megabytes Per 2400 Ft Reel[2] |
|---|---|---|---|
| 4096 | 0.96 | 28750 | 117.760 |
| 8192 | 1.61 | 17142 | 140.427 |
| 16384 | 2.92 | 9452 | 154.862 |
| 32768 | 5.54 | 4981 | 163.217 |
| 65536 | 10.79 | 2557 | 167.576 |

Notes:

[1] Includes 0.3 inch inter-record gap.

[2] 2,300 ft. useable recording length
(50 ft. leader, 50 ft. trailer)

used on computing systems whose arithmetic structure differs from that of the CDHF processing equipment.

There is little doubt that the most desirable format for archived data would use a universally recognized format such as ASCII character strings. However, as noted earlier, the already large quantity of reels of tape required to archive UARS data could increase to several tens of thousands of reels of CCTs if data are archived in the character format.

Further investigation of the question of the UARS archive medium and format will be essential to establish an archive which is user accessible, compatible with potential users' computing machinery and, at the same time, reasonable in its physical size.

An alternate solution to the archival storage medium is the use of optical disks. This technology is however still in its infancy and at the present time has a major weakness - error rates, which are much higher than magnetic tape. But because its advantages of more storage capacity in less space, faster access time and larger archival life are so promising, the implementation of optical disk systems for digital applications is a subject of considerable commercial and government research.

# 7.0 COMMUNICATIONS COSTS: CDHF/REMOTES

## 7.0 COMMUNICATIONS COSTS: CDHF/REMOTES

In order to derive cost estimates for CDHF/Remote communications a comparison was made for the following three modes:

- Packets
- Digital Service Leased Line
- Satellite Hop

The comparison was made under the following assumptions:

- Remote located 2000 miles from GSFC
- 12 Mbytes/day of traffic (average) between GSFC and a UARS Remote
- 38.5 Mbytes/day of traffic (average) between GSFC and an OPEN Remote.

The 12 Mbytes/day average between the UARS CDHF and a Remote is derived as follows: Total traffic from the UARS CDHF to the PIs per day is (see Table 3.1-1):

| | | |
|---|---|---|
| All of L3 | (Further analysis ) = | 48.35 |
| 10% of L0 | (Quicklook) = | 20.00 |
| 5% of L1 | (Dev. verif./anal.) = | 17.84 |
| 20% of L2 | (Dev. verif./anal.) = | 26.67 |
| | | 112.86 MB |

Additionally, the amount of data products transmitted from the PIs to the CDHF for the use of other investigators is assumed to be equivalent in volume to:

10% of L3 (Other users) = 4.82 MB.

Thus, this total two-way traffic is about:

112.86 + 4.83 = 117.69 MB/day.

Since there are ten instruments, the average amount of data associated with each instrument and hence the average two-way traffic per UARS Remote is:

117.69 / 10 = 12 MB/day.

The 12 MB/day derived here is just an average; traffic between the CDHF and a specific Remote can be more accurately estimated by using Table 3.1-1.

The 38.5 MB/day traffic between the OPEN CDHF and a Remote is similarly derived: In this case the sum of the following data volumes is taken:

All of L2 (Further analysis)
10% of L0 (Quicklook)
10% of L1 (Dev. verif./anal.)
Data equivalent in volume to 10% of L2 (Remote to CDHF for other users)

Using Table 3.2-1, this sum is equal to 1233 MB. Dividing by 32 instruments yields the number 38.5 MB.

Table 7.0-1 presents a summary of the communications costs to the Remotes. The cost figures for Packets and Digital Service Leased Line were derived from Fundamentals of Data Communications by Jerry Fitzgerald and Tom. S. Eason, 1978. The satellite communications costs were based on Planning Research Corporation (PRC) System Services Company's NASCOM Circuit Regression, which appears in Development of NASA DMS Performance/Cost Models, dated 5 January 1982. The details for the cost derivations are given in the paragraphs below.

TABLE 7.0-1

COMMUNICATIONS COSTS (DOLLARS/MONTH/REMOTE)

| Communication Mode | Costs (Dollars/Month/Remote) | |
|---|---|---|
| | UARS | OPEN |
| Packets | $18,768 | $59,274 |
| Leased Line | $2,139 | $2,139 |
| Satellite (Domestic) | $3,370 | $3,370 |
| Satellite (Overseas) | $19,430 | $19,430 |

## 7.1 Domestic

### 7.1.1 Packets

In addition to an installation fee of about $1,000, there are three cost factors:

- Packet Transmission Cost = $0.60/1000 packets
  $0.60/128,000 characters (bytes)
- Network Access Arrangement = $200/hour
- Network Interface Equipment = $400/month.

Note that packet transmission costs are independent of distance.

In the case of UARS, since a transmission of 12 MB requires about 2.78 hours on a 9.6 Kbps line, the monthly cost would be:

$$\frac{\$200}{\text{hr}} \times \frac{2.78\text{ hr}}{\text{day}} \times \frac{30\text{ day}}{\text{month}} + \frac{\$400}{\text{month}} + \frac{\$0.60}{128{,}000\text{ bytes}} \times$$

$$12{,}000{,}000\text{ bytes} \times 30\text{ days} = \$18{,}768/\text{month/UARS Remote}$$

An additional overhead cost must be added for packet header information.

For the OPEN case, since a transmission of 38.5 MB requires about 8.91 hours on a 9.6 Kbps line, the monthly cost would be:

$$\frac{\$200}{\text{hr}} \times \frac{8.91\text{ hr}}{\text{day}} \times \frac{30\text{ day}}{\text{month}} + \frac{\$400}{\text{month}} + \frac{\$0.60}{128{,}000\text{ bytes}} \times$$

$$38{,}500{,}000\text{ bytes} \times 30\text{ days} = \$59{,}274/\text{month/OPEN Remote}$$

### 7.1.2 Digital Service Leased Line

In addition to an installation fee of no more than $1,000 there are three cost factors:

- Intercity Line Mileage = $62/month + $0.93/mile (assuming 2000 miles)



C-2

- Service Terminals = \$134/month + \$1.34/mile (assuming 50 miles from end office)

- Network Interface Equipment = \$16/month

The total monthly cost would be:

$$\$62 + \$0.93 \times 2000 + \$134 + \$1.34 \times 50 + \$16 = \$2{,}139/\text{month/Remote, either UARS or OPEN}$$

Note that since this service would be provided on a 24 hours/day basis, considerably more than the assumed data volumes could be transmitted at the same cost.

### 7.1.3 Satellite Communications

The formula for computing the annual communications cost for a domestic satellite circuit is given by:

$$\text{FY 1982 K Dollars} = (11.98)[(\text{kilometers})(\text{Megabits})]^{0.3508}$$

with an error of ±7%.

Since 2,000 miles is about 3200 kilometers and 9,600 bits is about 0.01 megabits, the annual cost in kilodollars is:

$$\text{FY 1982 K Dollars} = (11.98)[(3{,}200)(0.01)]^{0.3508} = 40.41$$

The monthly recurring cost would thus be:

\$40,410 / 12 = \$3,370/month/Remote, either UARS or OPEN.

## 7.2 Overseas Circuits

The formula for computing the annual communications cost for an overseas satellite circuit is given by:

$$\text{FY 1982 K Dollars} = (700.6)(\text{Megabits})^{0.2389}$$

with an error of ±16%.

Since 9,600 bits is about 0.01 megabits, the annual cost in kilodollars is:

$$\text{FY 1982 K Dollars} = (700.6)(0.01)^{0.2389}$$
$$= 233.17$$

The monthly recurring costs would thus be:

$233,170 / 12 = $19,430/month/Remote, either UARS or OPEN.

# 8.0 POTENTIAL TECHNOLOGY APPLICATIONS

## 8.0 POTENTIAL TECHNOLOGY APPLICATIONS

The technology for implementing the UARS and OPEN data systems exists at the present time. However, there are technologies that should be available during the mission time-frames that could be utilized for a more cost effective or better performing system. The technologies examined are in the areas of data management, mass storage, software language development and communications.

### 8.1 Data Management

The most promising potentially applicable data management technology is that of the data base machine. Until recently data base management systems (DBMS's) have been software systems which executed on standard general purpose computers. However, two major limitations have surfaced under this implementation scheme. Data management systems that run on conventional computers run into bottlenecks when processing a large volume of transactions on very large (10 Gbytes) data bases. This is due to the data staging "bottleneck" between mass storage and main memory. The second limitation is that users are continually demanding more sophisticated DBMS capabilities such as backup and recovery, integrity and security controls, etc. These capabilities are needed by OPEN and UARS and require tremendous overhead. Consequently, a number of researchers have proposed the use of dedicated or specialized processors to execute data management functions. These are called data base machines (DM).

Several DM architectures are under investigation. All involve parallelism in one form or another and therefore take advantage of emerging VLSI technology. An example of a DM available today is a Britton-Lee computer designed specifically for DBMS processing. With software and

hardware the entire system can be purchased for about $200,000 (cost will vary depending on data base size and options). It is capable of data base access times equivalent to those obtained on a 5 to 10 MIPS standard computer with a software data base and there are plans to increase performance by another 5 to 10 fold. It will currently support up to 10 Gbytes of disk storage.

## 8.2 Mass Storage of Data

A common theme to the OPEN and UARS architectures discussed in this volume is that to achieve a balance between operational performance and system costs a hierarchy of computer memories/storage technologies is required. This hierarchy consists of a spectrum of cache/main memory, mass storage, and archival memory devices that span roughly six orders of magintude in both performance and cost. Because most technology involved in the existing memory hierarchy continues to reduce the per-bit storage cost at about the same rate, there will be no cross-over within the hierarchy within the near future. Therefore, memory hierarchies will continue to play a key role in the design of cost effective system architectures.

The storage technologies for accomplishing the objectives of the OPEN and UARS missions are well at hand. However, although there are numerous choices which can be made among alternate computer systems for performing production and communication tasks, there are only two choices for implementing the mass storage function. These choices, describe previously in this report are the IBM Mass Storage System (MSS) and the Masstor Virtual Storage System (VSS). Both these systems are basically automated magnetic tape-cartridge read/write systems that access the appropriate

cartridge, load it, and transfer the data to a staging disk in a matter of seconds. In the near term it does not appear that these devices will be supplanted. However, it can be anticipated that with the continuing price decrease in VLSI technology, more device intelligence will be built into mass storage devices. This would help remove the data-location burden from the CPU as well as minimize I/O traffic between mass storage and main memory. Additionally there could be an implementation in the mass-storage devices of such features as format initialization, limit checking, data compression and expansion, and error correction.

On the horizon the only apparent alternative to the magnetic cartridge mass storage devices seems to be the emerging optical disk storage systems. Optical disks promise a higher storage density and a lower per-bit cost than any other mass storage medium. Additionally they are they are made of materials that can be stored for many years without stringent environmental controls. However, opitcal disks suffer the drawback of being write-once devices. Although most magnetic tape is used in a write-once manner, there is a reluctance to utilize a new technology that forces this mode of operations.

At the present time, RCA has completed experimental optical disk systems that can record 5 Gbytes of data on one side of an optical disk at rates exceeding 100 Mbits/sec. These systems have provided a bit error rate of one-in-100 Mbits and can access any block of data in less than 0.5 seconds. There are plans to design a unit that would hold a number of optical disk platters that would be retireved and loaded as the need arose. It is planned that the worst case access time for a data block in this system would be about 5 seconds to retrieve data from a stored disk and .5

seconds if the disk were already on line. This type of system has been proposed to have 1.25 terabytes of storage.

Before optical data storage hardware becomes a reality, however, much work remains in the mechanics, the optics and the recording medium. Nonetheless, the current level of development activities suggests that operational systems will become widespread by the late 1980's or early 1990's.

A mass storage system that is exclusively optical disk does not appear feasible for OPEN and UARS-type projects because the write-once limitation could lead to a database size of over a terabyte. However, reversible data (Levels 0 and 1 for UARS and Level 0 for OPEN), which would rarely be altered, could be optically stored. An example of an advantage here could be the ease by which large quantities of this data could be recorded on a single disk (5 Gbytes or more) and sent by an express package service to the investigators. This could relieve a heavy I/O and communications burden from the CDHF.

## 8.3 Software Language Developments

The most likely major transition in languages that can be expected in the near future is the acceptance and use of the Ada language. Ada is currently under development by the Department of Defense (DOD) to be used in all of their software systems. Not only is it a powerful and flexible structured language, but it also serves as a program support environment, particularly for transportability, as well as supplying a methodology for life cycle software development, particularly in the area of configuration management. The use of Ada for OPEN and UARS would require massive

programmer retraining, the positive features of Ada may not outweigh this initial disadvantage. Moreover the cost benefit of Ada has not yet been proved.

Currently there are research efforts under way for producing compilers for automatic program generation in the sense that languages would be produced which would allow statements about <u>what</u> the program is to do to generate high-level language algorithms for stating <u>how</u> the program is to produce the desired results. For languages under current research, the compiler determines the sequence of procedures by analyzing the statements entered. This is in contrast to conventional languages in which control flow is built into the program itself.

At the present time there are no commercially available compilers for automatic program generation. It does not appear that one would be available for OPEN and UARS. Moreover, it remains to be seen whether increased hardware performance can overcome the potential slowness and inefficiency of multi-level compilers which first translate specifications into high-level languages and then into machine-language instructions.

## 8.4 Communications

Communications will be paced by advances in satellites and optical fibers.

In satellite communications, research in the areas of space diversity and time-division techniques, developments in antenna technology, sophisticated high-speed on-board switching, exploiting higher frequency sections of the spectrum, and on-board error detection and correction would provide for much broader wideband capabilities in space. However, under the communications traffic assumed by UARS, for example, recurring

satellite and terresterial communications costs were about the same. As long as rate structures are determined as a function of distance, this can be expected to remain the case. It remains to be seen if this policy will change.

Over the next few years local networks will be wire-based. If fiber is to compete, interfaces must be developed for fiber-optic systems that are compatible with coaxial networks such as Ethernet. Also, research is needed to define network topologies that best utilize fiber optics. Standards are now being established for defining a general class of terminal device for an optical fiber system. A goal would be the interchangability of the terminal device with a terminal device for a wire-based network.

Outside of local network applications, high-speed fiber optic buses may fill the need for fast parallel transfers between a mainframe and high-speed peripherals. This may serve to relieve any potential data staging bottlenecks between mass storage and main memory, as could be the case in the CDHF.

# APPENDIX A

# OPEN AND UARS MISSIONS

# ASSUMPTIONS AND INTERCOMPARISONS

# FLIGHT SEGMENT

<table>
<tr><th></th><th>UARS</th><th colspan="2">OPEN</th></tr>
<tr><td>Spacecrafts</td><td>1 Satellite</td><td colspan="2">4 Labs*<br>- IPL (Interplanetary Physics Lab)<br>- PPL (Polar Plasma Lab)<br>- GTL (Geomagnetic Tail Lab)<br>- EML (Equatorial Magnetosphere Lab)<br><br>* Contingency plan for a fifth lab (as spare)</td></tr>
<tr><td rowspan="7">Orbit</td><td rowspan="7">56° Inclination<br>600 Km Orbit</td><td>Min. Slant Range</td><td>Max. Slant Range</td></tr>
<tr><td>IPL Halo: 250 $R_E$</td><td>Halo: 250 $R_E$</td></tr>
<tr><td>PPL Initial: 300 to 3000 Km<br>Final: 300 to 3000 Km</td><td>Initial: 5 $R_E$<br>Final: 15 $R_E$</td></tr>
<tr><td>GTL 60 $R_E$</td><td>250 $R_E$</td></tr>
<tr><td>EML 2 $R_E$</td><td>12 $R_E$</td></tr>
<tr><td colspan="2">All labs have on-board propulsion systems to effect major alterations to orbital configurations</td></tr>
<tr><td colspan="2"></td></tr>
<tr><td>Lifetime (months)</td><td>18</td><td colspan="2">IPL: 36   GTL: 36   EML: 30   PPL: 24<br>- Simultaneous operations of 4 labs for 2 years (6 months min.)</td></tr>
</table>

FLIGHT SEGMENT (Concluded)

| | UARS | OPEN |
|---|---|---|
| Retrievable | Possibly | NO |
| Launch | STS<br>October 1988 | STS<br>IPL: February 1989<br>PPL: February 1990<br>GTL: February 1989<br>EML: August 1989 |
| Launch-Site | ETR | IPL: ETR<br>PPL: WTR<br>GTL: ETR<br>EML: ETR |
| On-Board Tape Recorders | YES | YES |
| Instrument Complement | Finalized<br>(See Appendix B) | Finalized<br>(See Appendix B) |
| On-Board Computer | YES | YES |
| • On-Board Data Processing | NONE | Some Data Reduction |
| • On-Board Data Compression | NONE | NONE |
| Design Goal for Unattended Operation | YES - 24 hours | Yes - 24 hours |

## SPACE/GROUND LINK

| | UARS | OPEN |
|---|---|---|
| Data Acquisition Communications | • TDRSS SSA (10 min. contact every orbit) (over 24 hour period)<br>• GSTDN back-up | • DSN (Deep Space Network) (Over 12 hour period)<br>• Occasional TDRSS Support |
| Telemetry Rates | • TR playback at 512 Kbps (data reversed), science plus engineering<br>• Real-time 32 Kbps<br>• Engineering only - 1 Kbps<br>• (1 Kbps to GSTDN-engineering only) | IPL / GTL / EML / PPL<br>Average (Kbps): 4 / 16 / 32 / 64<br>Playback (Kbps): 94 / 94 / 600 / 600 |
| Commanding Rates | • 1 Kbps nominal<br>• 125 bps emergency | Not Known |
| Telemetry Encoding | PCM | PCM |
| Telemetry Multiplex | TDM (firm) | TDM (Trade-off study to be made) |

## GROUND SEGMENT: PRINCIPAL INVESTIGATORS (PI)

| | UARS | OPEN |
|---|---|---|
| # of Remotes | 19 (Identical) | 42 (Not All Identical) |
| # Experiments | Total<br>10 + 9 TI's | IPL GTL EML PPL Total<br>7 5 9 11 32 + 5TI's |
| Location | Remote | Remote |
| PI Facility | • Funded by UARS project<br>• Minicomputer and necessary peripherals | • Funded by OPEN Project<br>• Minicomputer and necessary peripherals |
| Software Development | • To be done by PI for each instrument<br>• PI's will be provided software simulators of UARS/CDHF high speed processors<br>• Testing and integration to take place in UARS/CDHF | • To be done by PI for each instrument<br>• PI's will be provided software simulators of OPEN/CDHF high speed processors<br>• Testing and integration to take place in OPEN/CDHF |
| Communications with CDHF | Via 9.6 Kb hardwire line | Via 9.6 Kb hardwire line |
| Compatibility with CDHF | • Software<br>- similar operating system<br>- transportability of software<br>• Hardware<br>- same vendor | • Software<br>- similar operating system<br>- transportability of software<br>• Hardware<br>- same vendor |
| Facility Hardware | • Identical<br>• Minicomputers<br>• Graphics | • 7 mini's<br>• 21 Medi's<br>• 14 Maxi's<br>• All have graphics |

## GROUND SEGMENT: CENTRAL DATA HANDLING FACILITY (CDHF)

| | UARS | OPEN |
|---|---|---|
| Dedicated | YES | YES |
| Function | UARS data processing and data management | Open data processing and data management |
| Readiness | 3 months prior to launch | 3 months prior to launch |
| Raw Data Ingest Volumes | 1.6 X $10^9$ bits per day | About $10^{10}$ bits (all 4 labs) per day |
| Data Accessability | All PI's and TI's able to access data from all UARS instruments | All PI's, TI's and CI's able to access data from all OPEN instruments |
| Data Levels | Level 0: Raw data<br>Level 1: Calibrated data (reversible)<br>Level 2: Data converted to geophysical units (irreversible)<br>Level 3: Interpolation of geophysical parameters onto a common grid (irreversible) | Level 0: Raw data<br>Level 1: Calibrated data (reversible)<br>Level 2: Data converted to geophysical units (irreversible) |
| Routine Level Conversion of Data | Level 0 through Level 3 | • Level 0 to Level 1<br>• Level 1 to Level 2 at CDHF and remotes |
| Analysis | At remotes | In CDHF or remotes |
| Quick Look Capability | TBD | YES (within 8 hours) |
| Digital Communications with PI's | YES (9.6 Kb lines) | YES (9.6 Kb lines) |

GROUND SEGMENT: CENTRAL DATA HANDLING FACILITY (CDHF) (Continued)

| | UARS | OPEN |
|---|---|---|
| Operation | 2 shifts/day, 7 days a week (See Appendix C) | 3 shifts/day, 7 days a week (See Appendix C) |
| Operational Lifetime | 30 months (See Appendix C) | 48 Months (See Appendix C) |
| Availability Goal | $\geq$ 99% | $\geq$ 99% |
| Processing Capability | 2 MFLOPS/sec Min. Effective (est.) | 9 MFLOPS/sec Min. Effective (est.) |
| Maximum Mass Store on-line) Data Retention | 111 Gigabytes | 418 Gigabytes |
| Data Availability | • Level 0 data within 48 hours<br>• Level 1 data within 7 days<br>• Level 2 data within 10 days<br>• Level 3 data within TBD days | • Level 0 data available for quick-look within 8 hour after receipt of data at CDHF<br>• Level 1 data within 24 hours<br>• Level 2 data within 72 hours |
| Array Processing | YES | YES |
| Tracking/Orbit Computatation Support | YES | YES |
| Command Management | "Mailbox" | Pre-processing at CDHF (TBD) |
| Operation and Control Functions | by MSOCC<br>(Instrument related command requests will pass through CDHF) | by MSOCC<br>(Instrument related command requests will pass through CDHF) |

GROUND SEGMENT: CENTRAL DATA HANDLING FACILITY (CDHF) (Concluded)

| | UARS | OPEN |
|---|---|---|
| On-Line Retention Period | • Level 0: 10 days<br>• Level 1: 30 days<br>• Level 2: Life of mission + 1 year<br>• Level 3: Life of mission + 1 year | All Level 1 and 2 data for 100 days |
| Off-Line Storage | For all raw and processed data | • After 100 days to off-line store<br>• After 1 year to NSSDC |
| Off-Line Storage Requirement Over Mission Lifetime | 399 Gigabytes | 4,553 Gigabytes |
| System Redundancy | • No single point failure<br>(Common elements with OPEN/CDHF) | • No single point failure<br>(Common elements with UARS/CDHF) |
| Availability of Other Data Bases | Not Planned | Not Planned |

GROUND SEGMENT: DATA CAPTURE FACILITY (DCF)

| | UARS | OPEN |
|---|---|---|
| Dedicated | YES | YES (Colocated with CDHF) |
| Function | • Decommutates data to Level 0 experiment files<br>• Strip raw attitude<br>• Send playback engineering to POCC<br>• Quality check | • Decommutates data to Level 0 experiment files (TBD)<br>• Strip raw attitude<br>• Send playback engineering to POCC<br>• Quality check |
| Data Source | NASCOM/TDRSS | NASCOM/DSN, Occassional TDRSS support |
| Peak Input Rate | 512 Kbps (TBD) | 1 Mbps (TBD) |
| Output Rate | TBD | TBD |

# APPENDIX B

# UARS AND OPEN

# INSTRUMENT SELECTIONS

# UARS INSTRUMENT SUMMARY

| INVESTIGATION | INSTRUMENT | PI | INSTITUTION | LOCATION |
|---|---|---|---|---|
| **ENERGY INPUTS** | | | | |
| Energetic Particles | Particle Environment Monitor (PEM) | J.W. Winnigham | Southwest Research Institute | Dallas, TX |
| Ultraviolet Solar | Solar Ultraviolet Spectral Irradiance Monitor (SUSIM) | G.E. Brueckner | NRI | Washington, DC |
| Ultraviolet Solar | Solar Stellar Irradiance Comparison Monitor (SOLSTICE) | G.J. Rottman | U. of Colorado | Boulder, CO |
| **CHEMICAL SPECIES/TEMPERATURE** | | | | |
| Microwave Emission Radiometer | Microwave Limb Scanner (MLS) | J.W. Waters | Jet Propulsion Lab. | Pasadena, CA |
| Infrared Occulation Radiometer | Halogen Occulation Experiment (HALOE) | J.M. Russell, III | Langley Research Center | Hampton, VA |
| Infrared Emission Radiometer | Improved Stratospheric and Meso-spheric Sounder (ISAMS) | F. Taylor | Oxford | Oxford, England |
| Infrared Emission Radiometer (Cryogenic) | Cryogenic Limb Array Etalon Spectrometer (CLAES) | A.E. Roche | Lockheed | Palo Alto, CA |
| Solar Backscatter Ultraviolet Radiometer | Solar Backscatter Ultraviolet Experiment (SBUV) | Frederick | Goddard Space Flight Center | Greenbelt, MD |
| **WINDS** | | | | |
| Michelson Interferometer | Winds and Temperatures (WINTERS) | G. Thuillier | NCRS | Paris, France |
| Fabry-Perot Interferometer | High Resolution Doppler Imager (HRDI) | P.B. Hays | U. of Michigan | Ann Arbor, MI |

## UARS THEORETICAL INVESTIGATORS

| TI | Institution | Location |
|---|---|---|
| J.S. Chang | Lawrence Livermore Laboratory | Berkley, California |
| A. Gadd | United Kingdom Meteorological Office | London, England |
| D.M. Cunnold | Georgia Tech. | Atlanta, Georgia |
| M.A. Geller | Goddard Space Flight Center | Greenbelt, Maryland |
| W.J. Grove | Langley Research Center | Hampton, Virginia |
| J.R. Holton | Washington University | Seattle, Washington |
| A.J. Miller | NOAA Meteorological Center | Washington, D.C. |
| C.A. Reber | Goddard Space Flight Center | Greenbelt, Maryland |
| R.W. Zurek | Jet Propulsion Laboratory | Pasadena, California |

## OPEN INSTRUMENT SUMMARY

| INVESTIGATION | PI | INSTITUTION |
|---|---|---|
| **PPL** | | |
| Magnetic Fields | J.W. Winnigham | UCLA |
| Electric Fields | Mozer | UC, Berkeley |
| Plasma Waves | Shawhan | U of Iowa |
| Hot Plasma | Scudder | GSFC |
| Hot Plasma Composition | Shelley | Lockheed PARC |
| Cold Plasma | Chappell | MSFC |
| Energetic Particles | Higbie | Los Alamos SL |
| Energetic Particle Comp | Fritz | NOAA SEL |
| Auroral Imager VIS | Feldman | JHU |
| Auroral Imager UV | Torr | Utah State |
| Auroral Imager Xray | Imhof | Lockheed PARC |
| **EML** | | |
| Magnetic Fields | McPherron | UCLA |
| Electric Fields | Maynard (passive) | GSFC |
| | McIlwain (active) | UC, San Diego |
| Plasma Waves | Scarf | TRW |
| Hot Plasma | Parks | U of Washington |
| Hot Plasma Composition | Burch | Southwest RI |
| Cold Plasma | Chappell | MSFC |
| Energetic Particles | Higbie | Los Alamos SL |
| Energetic Particle Comp | Fritz | NOAA SEL |
| **GTL** | | |
| Magnetic Fields | Lepping | GSFC |
| Electric Fields | Mozer | UC, Berkeley |
| Plasma Waves | Gurnett | U of Iowa |
| Hot Plasma | Frank | U of Iowa |
| Energetic Particle Comp | Williams | NOAA SEL |
| **IPL** | | |
| Magnetic Fields | Behannon | GSFC |
| Plasma Waves | Kaiser | GSFC |
| Hot Plasma | Ogilvie | GSFC |
| Hot Plasma Composition | Gloeckler | U of Maryland |
| Energetic Particles | Lin | UC, Berkeley |
| Cosmic Rays | McDonald | GSFC |
| Gamma Rays | Teegarden | GSFC |

## ADDITIONAL INVESTIGATORS

| INVESTIGATION | INSTITUTION |
|---|---|
| **THEORY** | |
| M. Ashour-Abdalla | UCLA |
| M. Hudson | UC, Berkeley |
| D. Papadopoulos | U of Maryland |
| F. Rees | U of Alaska |
| C. Sonett | U of Arizona |
| **GROUND BASED** | |
| J. Dudeney | NERC/UK |
| R. Greenwald | JHU/APL |
| R. Vondrak | SRI |

# APPENDIX C

## UARS AND OPEN OPERATIONAL SCHEDULES

# UARS AND OPEN OPERATIONAL SCHEDULES (UPPER: UARS; LOWER: OPEN)

CDHF + Remote Terminal H/W in Place (−24) · Ground System Readiness (between −6 and 0) · UARS Launch (0) · UARS Termination (18) · Terminate CDHF Operation (30)

→ Months After UARS Launch

| −24 to −12 | −12 to 0 | 0 to 6 | 6 to 12 | 12 to 18 | 18 to 24 | 24 to 30 |
|---|---|---|---|---|---|---|
| 1 Shift | 2 Shifts | 2 Shifts | 2 Shifts | 2 Shifts | 1 or 2 Shifts | 1 or 2 Shifts |
| • CDHF/Remote Terminal Integration, PI Developed S/W transferred to CDHF, and S/W Testing in CDHF | • Simulation<br>• Testing | • Investigator Tasks<br>• Data Validation | • Investigator Tasks<br>• Routine Processing | • Investigator Tasks<br>• Routine Processing | – Reprocessing<br>– Support Data Analysis<br>– Backlog Processing (if needed) | – Reprocessing<br>– Support Data Analysis |

No Backlog (12)

CDHF + Remote Terminal H/W in Place (−24) · Ground System Readiness (between −6 and 0) · IPL Launch GTL Launch (0) · EML Launch (6) · PPL Launch (12) · All Flights Terminate (36) · Terminate CDHF Operation (48)

→ Months After First Open Launch

| −24 to −12 | −12 to 0 | 0 to 6 | 6 to 12 | 12 to 18 | 18 to 24 | 24 to 30 | 30 to 36 | 36 to 42 | 42 to 48 |
|---|---|---|---|---|---|---|---|---|---|
| 1 Shift | 2 Shifts | 2 Shifts | 3 Shifts | 3 Shifts | 3 Shifts | 3 Shifts | 3 Shifts | 1 or 2 Shifts | 1 or 2 Shifts |
| • CDHF/Remote Terminal Integration, PI Developed S/W transferred to CDHF, and S/W Testing in CDHF | • Simulation<br>• Testing | • Investigator Tasks<br>• Data Validation (IPL + GTL) | • Investigator Tasks<br>• Routine Processing (IPL, GTL)<br>• Routine Processing (IPL + GTL) and Data Validation (EML) | • Investigator Tasks<br>• Routine Processing (IPL, GTL, EML)<br>• Routine Processing (IPL, GTL, EML) and Data Validation (PPL) | • Investigator Tasks<br>• Routine Processing (All 4 Labs)<br>• Routine Processing (All 4 Labs) + Backlog Processing | • Investigator Tasks<br>• Routine Processing (All 4 Labs)<br>• Routine Processing + Backlog Processing | • Investigator Tasks<br>• Routine Processing (All 4 Labs)<br>• Routine Processing and Backlog Processing | – Reprocessing<br>– Support Data Analysis | – Reprocessing<br>– Support Data Analysis |

No Backlog of IPL + GTL (24) · No Backlog EML (30) · No Backlog PPL (36)

ORIGINAL PAGE IS OF POOR QUALITY

# BIBLIOGRAPHY

UARS


1. Preliminary Execution Phase Project Plan For Upper Atmosphere Research Satellite (UARS) (GSFC, May 1978).

2. Final Report of the Science Working Group (JPL Publication 78-54, July 1978).

3. Upper Atmosphere Research Satellite (UARS) Technical Report (GSFC, August 1979).

4. Upper Atmosphere Research Satellite (UARS) Conceptual Configuration Study of Candidate Instruments (GSFC, August 1979).

5. Upper Atmosphere Research Satellite (UARS) Data Handling Facilities (DHF) Management Plan (GSFC, August 1981).

6. UARS Instrument Processing Questionnaire response from the PI's, June-December 1981.

7. Upper Atmosphere Research Satellite (UARS) Principal Investigator Data Processing Requirements (CSC Publication CSC/TM-81/6241, September 1981).

8. UARS Ground Data Processing System Capability Document GSFC/565, 22 October 1981).

9. UARS Instrument Technical Report/Technical Descriptions, 1981.

10. UARS Instrument IRD's, 1981.


OPEN


1. Preliminary Execution Phase Project Plan For Origin of Plasmas in the Earth's Neighborhood (OPEN) (GSFC, September 1979).

2. Origin of Plasmas in the Earth's Neighborhood, Final Report of the Science Definition Working Group (GSFC, April 1979).

3. Data Information Systems Study For Origin of Plasmas in the Earth's Neighborhood (OAO Publication, July 1979).

4. Instrument Proposals, 1981.

**ENGINEERING & ECONOMICS RESEARCH, INC.**
1951 Kidwell Drive, Vienna, VA 22180
(703) 893-8600